```
OM Of: US-09-294-298-6 to: GenEmbl:*     out_format : pfs
Date: Jan 19, 2001  1:07 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09294298/runat.17012001_125501_27756/app_query.fasta_1.4359
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=0.500 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09294298_@CGN1_1_13912 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-294-298-6
Query length: 1325
Database: GenEmbl:*
Database sequences: 1118133
Database length: -173692196
Search time (sec): 10440.680000
```

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation …
gb_ro:AF050183	+	6950.00	2848.35	1.7e-150	4265	! AF050183 Rattus norvegicus GTP
gb_ro:AF058790	+	6912.50	2833.25	1.2e-149	4140	! AF058790 Rattus norvegicus Syn
gb_ro:AF058789	+	6858.00	2810.58	2.2e-148	4539	! AF058789 Rattus norvegicus Syn
gb_ro:AF048976	+	6292.50	2581.05	1.3e-135	4063	! AF048976 Rattus norvegicus mRN
gb_ro:AB016962	-	6280.50	2575.28	2.7e-135	4801	! AB016962 Rattus norvegicus mRN
gb_htg20:AL161903	+	4060.50	1653.62	6.0e-84	150956	! AL161903 Homo sapiens chromo
gb_pr7:HSDJ570F3	-	4059.50	1655.68	4.6e-84	94770	! AL050332 Human DNA sequence f
gb_htg20:AL161906	+	3987.00	1623.11	3.0e-82	169323	! AL161906 Homo sapiens chromo
gb_pr3:AF047711	+	2584.50	1071.76	1.5e-51	4368	! AF047711 Homo sapiens nGAP mRN
gb_htg7:AC019800	+	1257.50	518.21	1.0e-20	56366	! AC019800 Drosophila melanogas
gb_htg4:AC012162	+	1257.50	511.73	2.4e-20	191504	! AC012162 Drosophila melanoga
gb_htg4:AC012161	+	1257.50	511.43	2.5e-20	202741	! AC012161 Drosophila melanoga
gb_in1:AE003506	+	1257.50	509.33	3.3e-20	300994	! AE003506 Drosophila melanoga
gb_in1:AB011280	+	1173.50	498.06	1.4e-19	3984	! AB011280 Caenorhabditis elegan
gb_pr4:AK022662	+	1088.50	466.19	8.3e-18	2385	! AK022662 Homo sapiens cDNA FLJ
gb_htg6:AC016523	+	1082.00	445.96	1.1e-16	65961	! AC016523 Drosophila melanogas
gb_pr4:AK024488	+	1081.50	460.24	1.8e-17	4287	! AK024488 Homo sapiens mRNA for
gb_ro:AF053938	-	761.00	338.52	5.7e-10	829	! AF053938 Rattus norvegicus p135
gb_htg22:AL365274	-	710.00	289.50	1.2e-07	172027	! AL365274 Homo sapiens chromo
gb_in2:CEC07B5	+	676.50	283.64	1.2e-07	39660	! Z46266 Caenorhabditis elegans
gb_pr6:HS593C16	-	650.50	269.22	7.7e-07	81971	! AL035702 Human DNA sequence f
gb_htg12:AC027051	+	650.50	265.78	1.2e-06	156700	! AC027051 Homo sapiens chromo
gb_htg3:AC011492	+	613.00	250.49	8.5e-06	157803	! AC011492 Homo sapiens chromo
gb_htg17:AC073765	+	600.00	243.03	2.2e-05	237619	! AC073765 Mus musculus clone
gb_in1:AB011279	+	503.50	228.38	0.0001	2276	! AB011279 Caenorhabditis elegan
gb_htg3:AC011703	+	502.00	205.25	0.0028	159777	! AC011703 Drosophila melanoga
gb_pr1:AB011110	+	488.50	217.53	0.0006	5579	! AB011110 Homo sapiens mRNA for
gb_htg13:AC040900	+	479.50	196.24	0.0090	155332	! AC040900 Homo sapiens chromo
gb_htg20:AL161906	+	479.00	195.58	0.0098	169323	! AL161906 Homo sapiens chromo
gb_ro:RATGAPX	+	455.00	206.69	0.0024	3296	! L13151 Rat GTPase-activating p
gb_pr8:HUMGAPB	+	437.50	199.31	0.0061	3456	! M23612 Human GTPase-activating
gb_pr8:HUMGAPA	+	437.50	199.15	0.0070	4307	! M23379 Human GTPase-activating
gb_pl2:BGR272158	+	433.50	198.19	0.0070	3144	! AJ272158 Blumeria graminis f.
gb_om:BTGAPR	+	433.50	196.32	0.0089	4473	! X12602 Bovine mRNA for GTPase
gb_pat1:E13125	+	429.50	197.47	0.0077	2647	! E13125 Bovine gene for R-Ras f
gb_om:BTU30857	+	429.50	197.31	0.0078	2730	! U30857 Bos taurus R-Ras GTPase
gb_htg12:AC027051	+	417.00	170.76	0.2355	156700	! AC027051 Homo sapiens chromo
gb_in3:DME012609	+	416.00	190.49	0.0188	3502	! AJ012609 Drosophila melanogas
gb_in1:AB011283	+	395.50	190.25	0.0194	760	! AB011283 Caenorhabditis elegans
gb_pr4:AK026441	+	394.50	182.91	0.0496	2813	! AK026441 Homo sapiens cDNA: FL
gb_pr7:HSINSP4BP	+	381.00	177.37	0.1010	2837	! X89399 Homo sapiens mRNA for I

(handwritten annotation at left margin: "Chen", and "99.5%")

```
gb_ro:MMU20238       +   381.00   175.41   0.1298   4107  ! U20238 Mus musculus GTPase-
gb_in1:AB011285      +   378.50   182.84   0.0500    833  ! AB011285 Caenorhabditis ele?
gb_ro:AF086714       +   371.00   172.80   0.1814   3117  ! AF086714 Mus musculus rasc?
gb_ro:AF055883       +   360.50   182.79   0.0504    211  ! AF055883 Rattus norvegicus p

seq_name: gb_ro:AF050183

seq_documentation_block:
LOCUS       AF050183         4265 bp    mRNA            ROD      15-SEP-2000
DEFINITION  Rattus norvegicus GTPase activating protein SynGAP-c mRNA, complete
            cds.
ACCESSION   AF050183
VERSION     AF050183.2  GI:10140854
KEYWORDS    .
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4265)
AUTHORS     Kim,J.H., Liao,D., Lau,L.F. and Huganir,R.L.
TITLE       SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
            protein family
JOURNAL     Neuron 20 (4), 683-691 (1998)
MEDLINE     98240917
REFERENCE   2  (bases 1 to 4265)
AUTHORS     Kim,J.H. and Huganir,R.L.
TITLE       Direct Submission
JOURNAL     Submitted (20-FEB-1998) Neuroscience, Johns Hopkins Medical
            Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
            On Sep 15, 2000 this sequence version replaced gi:3044054.
COMMENT
FEATURES             Location/Qualifiers
     source          1..4265
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="hippocampus"
                     /dev_stage="adult"
     CDS             277..3792
                     /note="neuronal synaptic RasGAP; associates with the
                     PSD-95/SAP90 protein family; alternatively spliced
                     variants, SynGAP-a and SynGAP-b"
                     /codon_start=1
                     /product="GTPase activating protein SynGAP-c"
                     /protein_id="AAC40082.2"
                     /db_xref="GI:10140855"
                     /translation="MOSFKESHSHESLLSPSSAAEALELNLDEDSIIKPVHSSILGQE
                     FCFEVTTSSGTKCFACRSAAERDKWIENLDQRAVKPNKDNSRRVDNVLKLWIIAERELP
                     PKKRYCELCLDDMLYARTTSKPSRSASGDTVFWGEHFEFNNLPAVRALRLHYRDSDK
                     KRKKDKAGYVGLVTVPVATLAGRHETEQWYPVTLPTGSGSGMGSGGGSGGGSGG
                     KGKGGCPAVRLKARYQTMSILPMELYKEFAEYVTNYRMLCAVLEPALNVKGEEVAS
                     ALVHLQSTGKAKDFLSDMAMSEVDRFMERBHLIFRENTLATKAIEEYMRLIGQKYLK
                     DAIGEFIRALYESEENCEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPREL
                     KEVFASWRLRCAERGREDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTL
                     TLIAKVIQNLANFSKFTSKEDFLGFMNEFLELEWGSMQQFLYEISNLDTLTNSSSFEG
                     YIDLGRELSTLHALLWEVLPQLSKEALLKLGPLPRLLSDISTALRNPNIQRQPSRQSE
                     RARSQPMVLRGPSAEMOGYMMRDLNSSIDLOSFMARGLNSSMDMARLPSPTKEKPPP
                     PPGGGKDLFYVSRPPLARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGDGPGGRL
                     NSSSVSNLAAVGDLLHSSQASLTAALGLRPAPAGRLSQSGSSTTAAGMRLSOMGVTT
                     DGVPAQQLRIPLSFQNPLFHMAADGPPAGHGSGHGPPSSHHHHHHHRGGE
                     PGDTFAPFHGYSKSEDLSTGVPKPPAASILHSHSYSDEFGPSGTDFTRRQLSLQDNLQ
                     HMLSPPQITIGPQRPAPSGPGGSGGGGQPPPLQRGKSQQLTYSAAQKPRPS
                     SGNLLQSPEPSYGPARPRQQSLSKEGSIGGSGGSGGGGGLKPSITKQHSQTPSTLN
                     PTMPASERTVAWVSNMPHLSADIESAHIEREEYKLKEYSKSMDESRLDRVKEYEEIH
                     SLKERLHMSNRKLEEYERRLLSQEEQTSKILMQYVARLEQSEKRLRQQVEKDGIKS
                     IIGRLMLVEEELRRDHPAMAEPLPEPKKRLLDAQVERQLPLGPTNPRVTLAPPWNGL
                     APPAPPPPRLQITENGEFRNTADH"
BASE COUNT      936 a   1358 c   1203 g   768 t
ORIGIN

alignment_scores:
           Quality:  6950.00                 Length:   1328
           Ratio:       5.245                Gaps:        3
```

Percent Similarity: 99.774   Percent Identity: 99.699

alignment_block:
US-09-294-298-6 x AF050183

Align seg 1/1   to: AF050183   from: 1 to: 4265

```
    1 GlyGluThrGluLeuProGlnAlaProHisPhePrePheAlaProGlnGl  17
      |||||||||||||||||||||||||||||||||||||||||||||||||
   97 GGAGAGACTGAGCTGCCCAAGCACCCCATTTCCCATTTGCTCCCCAGCA  146

   17 nGlyPheLeuSerArgArgLeuLysSerSerIleLysArgThrLysSerG  34
      |||||||||||||||||||||||||||||||||||||||||||||||||
  147 AGGCTTCCTGAGCCGGAGGCTAAAAAGCTCCATCAAACGTACAAAGTCAC  196

   34 lnProLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPhe  50
      |||||||||||||||||||||||||||||||||||||||||||||||||
  197 AACCCAAACTTGACCGGACAAGCAGCTTTCGACAGATCCTGCCTCGCTTC  246

   51 ArgSerAlaAspHisAspArgAlaArgLeuMetGlnSerPheLysGluSe  67
      |||||||||||||||||||||||||||||||||||||||||||||||||
  247 CGAAGTGCTGACCATGACCGGGCCCGCTGATGCGAGCTTCAAGGAGTC  296

   67 rHisSerHisGluSerLeuLeuSerProSerSerAlaAlaGluAlaLeuG  84
      |||||||||||||||||||||||||||||||||||||||||||||||||
  297 TCACTCCCATGAGTCCCTGCTGAGTCCCAGCAGTGCTGCTGAGGCCCTGG  346

   84 luLeuAsnLeuAspGluLysProSerIleIleLeuLysProValHisSerIle  100
      |||||||||||||||||||||||||||||||||||||||||||||||||
  347 AGTCAACCTGGATGAGAGACTCCATTATCAAGCCAGTACCAGAGCTCCATC  396

  101 LeuGlyGlnLeuPheCysPheGluValThrThrSerGlyThrLysCys  117
      |||||||||||||||||||||||||||||||||||||||||||||||||
  397 CTGGGCCAGGAGTTCTGCTTTGAGGTAACAACATCGTCTCGGACAAAATG  446

  117 sPheAlaCysArgSerAlaAlaGluAlaArgAspLysTrpIleGluAsnLeuG  134
      |||||||||||||||||||||||||||||||||||||||||||||||||
  447 TTTTGCCGCAGTTCCGTCTGCAGCCCAAAGGGACAAATGGAATTGAGAATCTAC  496

  134 lnArgAlaValLysProAsnLysAspAsnSerArgArgValAspAsnVal  150
      |||||||||||||||||||||||||||||||||||||||||||||||||
  497 AGAGGGCGTGTGAAACCCAACAAGGACAACAGCCCGGGTAGATAACGTG  546

  151 LeuLysLeuTrpIleIleGluLeuAlaArgGluLeuProProLysLysArgTy  167
      |||||||||||||||||||||||||||||||||||||||||||||||||
  547 CTGAAACTATGGATCATAGAGCTCGAGAGCTGCCCCCAAGAAGCGATA  596

  167 rTyrCysLeuCysLeuAspAspSerPheMetLeuTyrAlaArgThrThrSerL  184
      |||||||||||||||||||||||||||||||||||||||||||||||||
  597 TTACTGCGAGTTATGCCTGGACGACATGCTCTATGCACGGACCACTTCCA  646

  184 ysProArgSerAlaSerGlyAspThrValPheThrValPheGluHisPheGlu  200
      |||||||||||||||||||||||||||||||||||||||||||||||||
  647 AGCCCCGCTCAGCGAGCAAGTCTTTTTGGGGCGAGCACTTCGAG  696

  201 PheAsnAsnLeuProAlaValAlaArgLeuAlaArgLeuHisLeuTyrArgAs  217
      |||||||||||||||||||||||||||||||||||||||||||||||||
  697 TTTAACAACCTGCCTGCCGTCGCGGCGCTGCCATCTGTACCGTGA  746

  217 pSerAspLysLysArgLysAspLysAspLysThrTyrValGlyTyrValTyrLeuValT  234
      |||||||||||||||||||||||||||||||||||||||||||||||||
  747 CTCGGACAAAAGCGGAAGAAGGACAAGGCAGGCTACGTTGGCCTGGTGA  796

  234 hrValProAlaAlaThrLeuAlaGlyArgHisPheThrGluThrTrpTyr  250
      |||||||||||||||||||||||||||||||||||||||||||||||||
  797 CTGTTCCAGTGGCCACCCTGGCTGGGCGCCACTTCACAGAGCAGTGGTAC  846

  251 ProValThrLeuProThrGlySerGlyGlySerGlyGlyMetGlySerGlyl  267
      |||||||||||||||||||||||||||||||||||||||||||||||||
  847 CCCGTGACCCTGCCAACAGGAAGTGGGGGTTCTGGGGTATGGGCTCGGG  896

  267 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyLysGlyGlyGlyC  284
      |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
  897 GGAGGAGGGGGGGGTCAGGGGGGCGGCTCAGGGGGCAAAGGGAAAGGAGGCT  946
      |||||||||||||||||||||||||||||||||||||||||||||||||
  284 ysProAlaValArgLeuLysAlaArgTyrGlnThrMetSerIleLeuPro  300

  947 GTCCTGCTGTGCGGCTGAAGGCCCGTGAAGGCCCGTTACCAGACAAATGAGTATCCTGCCC  996
      |||||||||||||||||||||||||||||||||||||||||||||||||
  300 MetGluLeuTyrLysGluPheAlaGluTyrValThrAsnHisTyrArgMe  317

 1046 ATGGAGCTATATAAGGAGTTGCAGAAATATGTGACCAACCACTACCGCAT  1046
      |||||||||||||||||||||||||||||||||||||||||||||||||
  317 tLeuCysAlaValLeuGluProAlaLeuAsnValLysGlyLysGluGluluV  334

 1096 GCTGTGTGCCGTGCTGGAGCCCGCCCTCAATGTCAAGGGCAAGGAGGAGG  1096
      |||||||||||||||||||||||||||||||||||||||||||||||||
  334 alAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLysAlaLysAsp  350

 1146 TCGCTAGTGCACTGGTTCACATCCTGCAAAGCACAGGCAAGGCAAGGAC  1146
      |||||||||||||||||||||||||||||||||||||||||||||||||
  350 alAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLysAlaLysAsp  367

 1196 PheLeuSerAspMetAlaMetSerGluValValAspArgPheMetGluArgGl  367
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1196 TTCCTTTCAGACATGGCCATGTCAGAGGTAGACCGGTTCATGGAGCGGGA  1196

 1246 uHisLeuIlePheArgGluAsnThrLeuAlaThrLysAlaIleGluGluluT  384
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1246 ACACCTCCATATTCCCGAGAACACGCTCGCCACTAAAGCCATAGAAGAGT  1246

 1296 yrMetArgLeuIleGlyGlnLysTyrLeuLysAspAlaIleAlaIleGlyGlyPhe  400
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1296 ATATGAGACTGATTGGCCAGAAATACCTTCAAGGATGCCATTGGGAGTTC  1296

 1346 IleArgAlaLeuTyrGluSerGluGluAsnCysGluValValAspProIleLeuLy  417
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1346 ATCCGGGCTCTGTATGAAATCTGAGGAGAACTGTGAAGTAGACCCCATCAA  1346

 1396 sCysThrAlaSerSerLeuAlaGluHisGlnAlaAsnLeuArgMetCysC  434
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1396 GTGCACAGCGTCCAGTCTGGCAGAGCACCAGGCCAACCTGCGGATGTGCT  1396

 1446 ysGluLeuAlaLeuCysLysLysValValAsnSerHisCysValPheProArg  450
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1446 GTGAGTTGGCCCTGTGCAAGTGGTCAACTCCCATTGCGTGTTCCCGAGG  1446

 1496 GluLeuLysGluValAlPheAlaSerTrpArgLeuArgCysAlaGluArgGl  467
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1496 GAGCTGAAGGAGGTGTTTGCATCATGCGGCTGCTGTGCAGAGCGGGG  1496

 1546 yArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuArgP  484
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1546 CCGGGAGGACATTGCTGACCAGGCTGATCAGCGGCGCCTCCTGCTGCGCCT  1546

 1596 heLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGlnGlu  500
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1596 TCCTCTGCCCGGCCAGCCATCATGTCGCCCAGTCTGTTTGGACTGATGCAGGAG  1596

 1646 TyrProAspGlnThrSerArgThrLeuThrLeuAlaLysValIll  517
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1646 TACCCAGATGAGCAGACCTCACGAACCCTCACCCTCGCCAAGGTTAT  1646

 1696 eGlnAsnLeuAlaAsnPheSerLysGluAspPheLeuGluG  534
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1696 CCAGAACCTGGCCAACTTTTCCAAGTTTACCTCAAAGGAGGACTTCCTGG  1696

 1746 lyPheMetAsnGluPheLeuGluLeuGluTrpGlySerMetGlnGlnPhe  550
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1746 GCTTCATGAACGAGTTTCTGGAGCTGGAATGGGGTTCTATGCAGCAATTC  1746

 1796 LeuTyrGluIleSerAsnLeuAspThrLeuThrArgAsnSerSerPheGl  567
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1796 TTGTATGAGATATCCAACCTGGACACTGACCAACACGAGCAGTTTTGA  1796

 1846 uGlyTyrIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuT  584
      |||||||||||||||||||||||||||||||||||||||||||||||||
 1846 GGGCTACATAGACTTGGGGCCCGAGCTCTCCACACTTCACGCCCTGCTCT
```

```
 584 rpGluValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGlyPro  600
     |||||||||||||||||||||||||||||||||||||||||||||
1847 GGGAGGTGCTGCCCAGCTCAGCAAGGAGAGCCCTCCTGAAGCTGGGGCCG 1896

 601 LeuProArgLeuLeuSerAspIleSerThrAlaLeuArgAsnProAsnIl  617
     |||||||||||||||||||||||||||||||||||||||
1897 CTGCCCCGGCTCCTCAGCGACATCAGCACAGCCCTGAGGAACCCTAACAT 1946

 617 eGlnArgGlnProSerArgGlnSerGluArgAlaArgSerGlnProMetV  634
     ||||||||||||||||||||||||||||||||||||||||
1947 CCAAAGGCAGCCGAGCCGCCAGAGCAGGCGCTCGGTCTCAGCCCCATGG  1996

 634 alLeuArgGlyProSerAlaGluMetGlnGlyTyrMetMetArgAspLeu  650
     |||||||||||||||||||||||||||||||||||||||
1997 TGCTGCGCGGCCCTCAGCCGAAATGCAGGGCTACATGATGCCGGGACCTC 2046

 651 AsnSerSerIleAspLeuGlnSerPheMetAlaArgGlyLeuAsnSerSe  667
     ||||||||||||||||||||||||||||||||
2047 AACAGCTCCATCGACCTTCAGTCCTTCATGGCTCGAGGCCTCAACAGCTC 2096

 667 rMetAspMetAlaArgLeuProSerProThrLysGluLysProProProP  684
     |||||||||||||||||||||||||||||||||
2097 TATGGACATGGCTCGCCTCCCCTCCCCAACCAAGGAGAAACCCCGCCGC  2146

 684 roProProGlyGlyLysAspLeuPheTyrValSerArgProProLeu     700
     ||||||||||||||||||||||||||||||||
2147 CCCTCCCCGGTGGGGTAAAGACCTGTTCTATGTGAGCCGGCCACCACTG  2196

 701 AlaArgSerProAlaTyrCysThrSerSerSerAspIleThrGluPr     717
     ||||||||||||||||||||||||||||||
2197 GCCCGGTCCTCCCCAGTATCGACGAGCAGCTCGGACATCACAGAGCC    2246

 717 oGluGlnLysMetLeuSerValAsnLysSerValSerMetLeuAspLeuG  734
     ||||||||||||||||||||||||||||
2247 GGAGCAGAAGATGCTGAGTGTCAACAAGAGTGTGTCCATGCTGGACCTGC 2296

 734 lnGlyAspGlyProGlyGlyArgLeuAsnSerSerSerValSerAsnLeu  750
     |||||||||||||||||||||||||||||||
2297 AGGGCGACGGGCCGGGCGGCCGCCTTAACAGCAGTAGTGTTTCCAACCTG 2346

 751 AlaAlaValGlyHisArgAspLeuHisSerSerGlnAlaSerThrAlaAl  767
     ||||||||||||||||||||||||||||||
2347 GCAGCTGTTGGGGACCGTTGCACTCAAGCCAGGCTTCACTGACAGCAGC  2396

 767 aLeuGlyLeuArgProAlaProAlaGlyArgLeuArgLeuSerGlySerG  784
     ||||||||||||||||||||||||||||||
2397 CTTGGGGTTGCGGCCTGCACCTGCCGGGGCCCTCCCCAAGGGAGTGGCT  2446

 784 erSerIleThrAlaAlaGlyMetArgGlyLeuSerGlnMetGlyValThrThr 800
     ||||||||||||||||||||||||||||
2447 TCTCTTCCATATGGCTGCCGATGGCGCCCTCAGCCAGATGGGTGCACTACG 2496

 801 AspGlyValProAlaGlnLeuGlnLeuArgIleProLeuSerPheGlnAsnPr 817
     |||||||||||||||||||||||||||||
2497 GATGGTGTCCCCGCCCAGCAACTGCCCATCCCCTCTTTCCTTCCAGAAACCC 2546

 817 oLeuPheHisMetAlaAlaAspGlyProGlyProAlaGlyHisGlyG     834
     |||||||||||||||||||||||||||||
2547 TCTCTTCCATATGGCTCCGATGGACCCAGGCCCCCAGGCCATGGAG     2596

 834 lySerSerGlyHisGlyProProSerHisHisHisHisHisHis        850
     ||||||||||||||||||||||||||||
2597 GGAGCAGTGGCCATGGGGCCTCCCATCCCCATCACCACCACCAT        2646

 851 HisHisHisArgGlyGlyArgProProGlyProGlyAspThrPheAlaProPheHi 867
     ||||||||||||||||||||||||||||
2647 CACCATCACCGAGGGGAGAACCCCAGGGGAACACTTTTGCCCGTTCCA     2696

 867 sGlyTyrLysSerGluAspGlyLeuSerThrGlyValProLysProProA  884
     |||||||||||||||||||||||||||
2697 TGGCTATAGCAAGAGCGAGGACCCTCTACAGGGGGTCCCTAAGCCCCCTG 2746
```

```
 884 lAlaAlaSerIleLeuHisSerHisSerTyrSerAspGluPheGlyProSer 900
     ||||||||||||||||||||||||||||||||||
2747 CGGCCTCCATCCTTCACAGCCCAGCTACAGTGATGAGTTTGACCCTCT    2796

 901 GlyThrAspPheThrArgArgGlnLeuSerLeuGlnAspAsnLeuGlnHi  917
     ||||||||||||||||||||||||||||||||
2797 GGTACTGATTTTACCCGTCGGCAGCTCCACTTCAGGACAACCTACAGCA  2846

 917 sMetLeuSerProProGlnIleThrIleGlyTyrProGlnArgProAlaProS 934
     |||||||||||||||||||||||||||||||||
2847 CATGCTCTCCCCGCCCAGATCACCATCGGTCCCCAGAGGCCAGCTCCCT  2896

 934 erGlyProGlyGlyGlySerGlyGlyGlyGlyGlyGly             950
     ||||||||||||||||||||||||||||||||
2897 CAGGGCCAGGAGGGGGCAGTGGTGGGGGCAGTGTGGGGGCCGGTGGGGGC 2946

 951 GlnProProProLeuGlnArgGlyLysSerGlnGlnLeuThrValSerAl  967
     ||||||||||||||||||||||||||||||||||
2947 CAGCCACCTCCCTTGCAGAGGGGCAAAATCACGCAGTTGACAGTGAGTGC 2996

 967 aAlaGlnLysProArgProSerGlyAsnLeuLeuGlnLeuSerProGluP  984
     ||||||||||||||||||||||||||||||||
2997 TGCCCAGAAAACCCGCCCTCCAGCGGCAACCTATTGCAGTCCCCGGAAC  3046

 984 roSerTyrGlyProArgProArgGlnLeuSerLeuSerLysGluGly    1000
     |||||||||||||||||||||||||||||||||||
3047 CAAGTTATNGTCCTGCCCGTCCACGGCAACGAGCCTCAGCAAAGAGGGC  3096

1001 SerIleGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyLeuLy       1017
     ||||||||||||||||||||||||||||||||
3097 AGCATTGGGGGGCACGGGGCACGGCGTGCGGAGGGGGTGGGGGGCTCAA  3146

1017 sProSerIleThrLysGlnHisSerGlnThrProSerThrLeuAsnProT 1034
     ||||||||||||||||||||||||||||||||||
3147 GCCCTCCATCACCAAGCAGCAGCATTCCCAGACTCCATCCACGCTGAACCCCA 3196

1034 hrMetProAlaSerGluArgThrValAlaTrpValSerArgAsnMetProHis 1050
     ||||||||||||||||||||||||||||||
3197 CGATGCCCGCCTCGAGCCGACTGTAGCCTGGGTGTCCAATATGCCTCAC  3246

1051 LeuSerAlaAspIleGluSerAlaHisIleGluArgGluGluTyrLysLe 1067
     ||||||||||||||||||||||||||||||||||
3247 CTGTCCGCTGCACATCGAGAGTGCACACATTGAGCGGGAAGAGTACAAGCT 3296

1067 uLysGluTyrSerLysSerMetAspGluSerArgLeuAsnProArgValLysG 1084
     ||||||||||||||||||||||||||||||||
3297 GAAGGAGTACTCGAAGTCCATGGACGAGAGCCGACTGGACAGGGTGAAGG 3346

1084 luTyrGluGluGlnIleHisSerLeuLysGluArgLeuHisMetSerAsn 1100
     ||||||||||||||||||||||||||||
3347 AGTACGAGGAGGAGATCCACTCACTGAAAGAAAGGCTACACATGTCCAAC 3396

1101 ArgLysLeuGluGluValGluGlnArgArgLeuSerGlnGluGlnGlnTh 1117
     ||||||||||||||||||||||||||||||
3397 CGGAAGCTGGAAGAGTACGAGCGGGAGCGGCTGTCTCCCAGGAAGCAGAC 3446

1117 rSerLysIleLeuMetGlnTyrGlnAlaArgLeuGluGlnSerGluAlaArgLeuLysA 1134
     ||||||||||||||||||||||||||||||
3447 CAGCAAGATCCTGATGCAGTACCAAGCCCGCCTGGAGCAGAGCGAGAAGC 3496

1134 rgLeuArgGlnGlnGlnValGluLysGlnAspSerGlnIleLeuLysSerIleIle 1150
     ||||||||||||||||||||||||||||||
3497 GCTTGAGGCAGCAGCAGGTGGAGAAGCAGGACTCCCAGATCAAGAGCATCATT 3546

1151 GlyArgLeuMetLeuValGluGluValGluArgArgAspHisProAlaMe 1167
     ||||||||||||||||||||||||||||||
3547 GGCAGGCTGATGCTGGTGGAGGAGGAGCTGCGCCGGGACCACCCCGCCAT 3596

1167 tAlaGluProLeuProGluProLysLysLysArgLeuLeuAspAla.Gln... 1182
     |||||||||||||||||||||| :::
3597 GGCTGAGCCGCTGCCTGAACCCAAGAAGAGGCTGCTGGACGCGTCAGGTGG 3646

1183 ArgGlySerPheProProTrpProValGlnThrArgVal...ArgTrpPr 1198
```

```
                 ||||||||||||||||||||||||||||||||||||||||||    ||||||||
3647  AGAGGCAGCTTCCCCCTTGGGCTCCAACAAACCGCCTGTGACGCTGGCC 3696

1198  oHisLeuGlyThrAlaTrpProProGlnProHisProProGlyC 1215
                 ||||||||||||||||||||||||||||||||||||||
3697  CCACCTTGGAACGGCCTGGCCCCCAGCCGCCCCCCCACCCCCGGCT 3746

1215  ysArgSerGlnArgThrAlaSerSerGlyThrProGlnThrThrSerPro 1231
                 ||||||||||||||||||||||||||||||||||||||||
3747  GCAGATCACAGAGAACGGCGAGTTCCGAACACCGCAGACCACTAGCCCA 3796

1232  ProSerIleThrAspLeuLeuProCysAlaProTyrProGlyProProSe 1248
                 ||||||||||||||||||||||||||||||||||||||||
3797  CCCAGCATCACAGACCTCCTTCCCTGTGCACCCTACCCCGGCCCACCCAG 3846

1248  rValThrAspLeuProSerAlaProAspArgProGlyThrSerProThrT 1265
                 ||||||||||||||||||||||||||||||||||||||||
3847  CGTCACAGACCTCCTTCCCAGTGCACCCGACCCTGGAACATCACCAACCA 3896

1265  hrArgThrGlyArgHisGlnGlyThrAlaGlyLeuSerProLeuThrPro 1281
                 ||||||||||||||||||||||||||||||||||||||||
3897  CCAGGACTGGACGTCACCCAAGGGACAGCGGATTGTCTCCCTTAACGCCT 3946

1282  ProTrpGlyThrHisLeuSerThrProLeuLeuHisSerArgArgGluSe 1298
                 ||||||||||||||||||||||||||||||||||||||||
3947  CCTTGGGGACCCACCATCTGTCAACCCCCACTGCTCCATTCCAGGAGGAGAG 3996

1298  rGlyThrLeuSerCysProLeuThrProGlyHisHisLeuProHisThrA 1315
                 ||||||||||||||||||||||||||||||||||||||||
3997  TGGGACCCTCAGCTGCCCTCTCACCCCCAGGACACCACTACCCCACACAG 4046

1315  spProPheThrLeuGlyCysTyrProHisPro 1325
                 |||||||||||||||||||||||||||||
4047  ACCCCTTCACTCTGGGGTGCTATCCCCATCCT 4078
```

seq_name: gb_ro:AF058790

seq_documentation_block:
```
LOCUS       AF058790    4140 bp    mRNA    ROD    09-OCT-1998
DEFINITION  Rattus norvegicus SynGAP-b mRNA, complete cds.
ACCESSION   AF058790
VERSION     AF058790.1  GI:3722228
KEYWORDS    .
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4140)
  AUTHORS   Kim,J.H., Liao,D., Lau,L.F. and Huganir,R.L.
  TITLE     SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
            protein family
  JOURNAL   Neuron 20 (4), 683-691 (1998)
  MEDLINE   98240917
REFERENCE   2  (bases 1 to 4140)
  AUTHORS   Kim,J.H. and Huganir,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-APR-1998) Neuroscience, Johns Hopkins Medical
            Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
REFERENCE   3  (bases 1 to 4140)
  AUTHORS   Kim,J.H. and Huganir,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-1998) Neuroscience, Johns Hopkins Medical
            Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
  REMARK    Sequence update by submitter
COMMENT     On Oct 9, 1998 this sequence version replaced gi:3065890.
FEATURES             Location/Qualifiers
     source          1..4140
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="hippocampus"
                     1..3750
                     /note="neuronal RasGAP; contains PH domain, C2 domain, GAP
```

domain in the N-terminal region, and a T/SXV motif at the C-terminus"
```
                     /codon_start=1
                     /product="SynGAP-b"
                     /protein_id="AAC63511.1"
                     /db_xref="GI:3722229"
                     /translation="MGLRPPTPTPSGGSGGSGSLPPPSHRQPLRRRCSSCCFPGEYHLG
                     RSRRKSVPGGKOYSMEAAPAAPFRPSQGFLSRRLKSSIKRTKSQPKLDRTSSFRQILP
                     RFRSADHDRARLMQSFKESHSHESLLSPSSAAEALELNLDEDSIIKPVHSSIIGQEFC
                     FEVTTSSGTKCFACRSAAERDKWIENLQRAVKPKDNSRRVDNVLKLWIIEARELPPK
                     KRYYCELCLDDMLYARTTSKPRSASGDTVFWGEHFEFNNLPAVRALRLHLYRDSDKKR
                     KKDKAGYVGLVTVPVAFLAGRHFTEQWYPVTLPTGSGSGGMGSGGGSGGSGGSGGKG
                     KGGCPAVRLKARYQTMSLPMELYKEFAEYVTNHYRMLCAVLEPALNVKGKEEVASAL
                     VHILQSTGKAKDFLSDMAMSEVDRFMEREHLIFRENTLATKAIEEYMRLIGQKYLKDA
                     IGEFIRALYESEENCEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKE
                     VFASWRLRCAERGREDIADRLISASLFRLCPAIMSPSLFGLMQEYPDEQTSRTLTL
                     IAKVIQNLANFSKFTSKEDFLGFMNEFLELEWGSMQQFLYEISNLDTLTNSSSFEGYI
                     DLGRELSTLHALLWEVLPQLSKEALLKLGPLPRLLSDISTALRNPNIQRQPSOSERA
                     RSQPMVLRGPSAEMQGYMMRDLNSSIDLQSFMARGLNSSMDMARLPSPTKEKPPPPP
                     GGGKDLFYVSRPPLARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGDGPGGRLNS
                     SSVSNLAAVGDLLHSSQASLTAALGLRPAPAGRLSQSGSSTTAAGMRLSQMGVTTDG
                     VPAQQLRIPLSFQNPLFHMAADGPGPPAGHGGSSHGPPSSHHHHHHHRRGEPPG
                     DTFAPFHGYSKSEDLSTGVPKPPAASIHSHSYSDEFGPSGTDFTRRQLSLQDNLQHM
                     LSPPQITIGPORPAPSGPGSGSGGGCQOPPLQRGKSQOLTVSAAQKPRPSSG
                     NLLQSPEPSYGPARPRQQSLSKEGSIGGSGSGGGGLKPSITKQHSQTPSTLNPT
                     MPASERTVAWVSNMPHLSADIPSAHIEREEYLKEYSKSMDESRLDRVKEYEEIHSL
                     KERLHMSNRKLEEYERRLLSQEEQTSKILMQYQARLEQSEKRLRQQQVEKDSQIKSII
                     GRLMLVEEELRRDHPAMAEPLPEPKKRLLDAQRGSFPPWVQQTRV"
```

```
BASE COUNT      902 a   1328 c   1182 g    728 t
ORIGIN
```

alignment_scores:
```
     Quality:  6912.50                    Length: 1319
     Ratio:    5.257                       Gaps:   1
Percent Similarity: 99.697     Percent Identity: 99.621
```

alignment_block:
US-09-294-298-6 x AF058790

Align seg 1/1 to: AF058790   from: 1 to: 4140 ..

```
  8  AlaProHisPhePheProAlaProGlnGlnGlyPheLeuSerArgArgLe  24
     ||||||  ||||:||||||||||||||||||||||||||||||||||
184  GCCCCCGCCGCCCTTCCGGCCCCTCCGCAGGGCTTCCTGAGCCGGAGGCT 233

 24  uLysSerSerIleLysArgThrLysSerGlnProLysLeuAspArgThrS  41
     ||||||||||||||||||||||||||||||||||||||||||||||||
234  AAAAAGTCCATCAAAACTACAAAGTCACAACCCAAACTTGACCGGACCA  283

 41  erSerPheArgGlnIleLeuProArgPheArgSerAlaAspHisAspArg  57
     |||||||||||||||||||||||||||||||||||||||||||||||||
284  GCAGCTTTCGACAGATCCTGCCTCGCTTCCGAAGTGCTGACCATGACCGG 333

 58  AlaArgLeuMetGlnSerPheLysGluSerHisSerHisGluSerLeuLe  74
     |||||||||||||||||||||||||||||||||||||||||||||||
334  GCCCGGCTGATGCAGAGCTTCAAGGAGTCTCACTCCCATGAGTCCTGCT  383

 74  uSerProSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAspS  91
     |||||||||||||||||||||||||||||||||||||||||||||
384  GAGTCCCAGCTGCTGCGAGCCCTGAGCTCAACCTGGATGAAGACT  433

 91  erIleIleLysProValHisSerSerIleLeuGlyGlnGluPheCysPhe 107
     ||||||||||||||||||||||||||||||||||||||||||||||||||
434  CCATTATCAAGCCAGTACACAGCTCCATCCTGGGCCAGGAGTTCTGCTTT 483

108  GluValThrThrSerSerGlyThrLysCysPheAlaCysSerArgAlaAl 124
     ||||||||||||||||||||||||||||||||||||||||||||||||||
484  GAGGTAACAACATCGTCTGGGACAAAATGTTTTGCCTGTCGGTCTGCAGC 533

124  aGluArgAspLysTrpIleGluAsnLeuGlnArgAlaValLysProAsnL 141
     |||||||||||||||||||||||||||||||||||||||||||||||
534  CGAAAGGGACAAATGATTGAGAATCTACAGAGGGCTGTGAAACCCAACA  583
```

CDS

```
141  ysAspAsnSerArgArgValAspAsnValLeuLysLeuTrpIleIleGlu  157        441  alValAlaAsnSerHisCysValPheProArgGluLeuLysGluValPheAla  457
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
584  AGGACAACACGCCGCGGGTAGATAACGTGCTGAAACTATGGATCATAGAA  633       1484  TGGTCAACTCCCATTGCGTGTTCCCGAGGGAGCTGAAGGAGGTGTTTGCA  1533

158  AlaArgGluLeuProProLysLysArgTyrTyrCysGluLeuCysLeuAs  174        458  SerTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAspAr  474
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
634  GCTCGAGAGCTGCCCCCCAAGAAGAGCGATTACTGCGAGTTATGCCTGGA  683       1534  TCATGGCGGCTGCGCTGTGCAGAGCGCGGGGAGGACATTGCTGACAG  1583

174  pAspMetLeuTyrAlaArgThrThrSerLysProArgSerAlaSerGlyA  191        474  gLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetS  491
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
684  CGACATGCTCTATGCACGGACCACTTCCAAGCCCCGCTCAGCCTCAGGAG  733       1584  GCTGATCAGCGCCTCGCTCTTCCTGCGCTTCCTCTGCCCGGCCATCATGT  1633

191  spThrValPheTrpGlyGluHisPheGluPheAsnLeuProAlaVal     207        491  erProSerLeuPheGlyLeuMetGlnThrTyrProAspGluGlnThrSer  507
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
734  ACACTGTCTTTTGGGCGGAGCACTTCGAGTTTAACAACCTGCCTGCTGTC  783       1634  CGCCCAGTCTGTTTGGACTCATGCAGGAGTACCCCAGATGAGCAGACCTCA  1683

208  ArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysArgGlyLy  224        508  ArgThrLeuThrLeuIleAlaLeuAlaLysValIleGlnAsnLeuAlaAsnPheSe  524
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
784  CGGGCCCTGCGGCTGCATCTGTACCGTGACTCGGACAAAAAGCGGAAGAA  833       1684  CGAACCCTCACCCTCATCGCCAAGGTTATCCAGAACCTGGCCAACTTTC  1733

224  sAspLysAlaGlyTyrValGlyLeuValThrValProValAlaThrLeuA  241        524  rLysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuG  541
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
834  GGACAAGGCAGGCTACGTTGGCCTGGTGACTGTTCCAGTGGCCACCCTGG  883       1734  CAAGTTTACCTCAAAGGAGGACTTCCTGGGCTTCATGAACGAGTTTCTGG  1783

241  laGlyArgHisPheThrGluGlnTrpTyrProValThrProThrGly     257        541  luLeuGluTrpGlySerMetGlnInPheLeuTyrGluIleSerAsnLeu  557
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
884  CTGGGCGCCACTTCACAGAGCAGTGGTACCCCGTGACCCCTGCCAACAGGA  933       1784  AGCTGGAATGGGGTTCTATGCAGCAATTCTTGTATGAGATATCCAACCTG  1833

258  SerGlyGlySerGlyGlyMetGlySerGlyGlyGlyGlyGlySerGlyGl  274        558  AspThrLeuThrAsnSerSerSerPheGluGlyTyrIleAspLeuGlyAr  574
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
934  AGTGGGGGGCTCTGGGGGTATGGCCTCGGGGGAGGAGGGGGTCAGGGGGG  983       1834  GACAACACTGACCAACAGCAGCAGTTTGAGGGCTACATAGACTTGGGCCG  1883

274  yGlySerGlyGlyLysGlyLysGlyGlyCysProAlaValArgLeuLysA  291        574  gGluLeuSerThrLeuHisIleAlaLeuLeuTrpGluValLeuProGlnLeuS  591
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
984  CGGCTCAGGGGGCAAAAGGAAAGGAGGCTGTCCTGCTGTGCGGGCTGAAGG  1033      1884  CGAGCTCTCCACACTTCACGCCCTGCTCTGGGAGGTGCTGCCCCAGCTCA  1933

291  laArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLysGluPhe  307        591  erLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSerArgAsp  607
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1034 CCCGTTACCAGACAATGAGTATCCTGCCCATGGAGCTATATAAGGAGTTT  1083      1934  GCAAGGAAGCCCTCCTGAAGCTGGGCCCGCTGCCCCGGCTCCTCAGCGAC  1983

308  AlaGluTyrValThrArgAsnHisTyrArgMetLeuCysAlaValLeuGluPr  324        608  IleSerThrLeuArgAsnProAsnIleGlnArgGlnProSerArgGl  624
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1084 GCAGAATATGTGACCAGAAACCACTACCGCATGCTGTGTGCCGTGCTGGAGCC  1133      1984  ATCAGCACCAGCCCTGAGGAACCCTAACATCCAAAGGCAGCCGAGCCGCCA  2033

324  oAlaLeuAsnValLysGlyLysGluLysAlaSerAlaLeuValHisI  341        624  nSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSerAlaG  641
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1134 CGCCCTCAATGTCAAGGGCAAGGAGAGGGAGTCGCTAGTGCACTGGTTCACA  1183      2034  GAGCGAGCGCCTCCGGTCGTCTCAGCCCATGGTGCTGCGGGGGCCGTCAGCCCG  2083

341  leLeuGlnSerThrGlyLysLysAlaLysAspPheLeuSerAspMetAlaMet  357        641  luMetGlnLeuTyrMetArgAspLeuAsnSerSerIleAspLeuGln  657
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1184 TCCTGCAAAGCCACAGGCCAAGGACTTCCTTTCAGACATGGCCATG  1233      2084  AAATGCAGCTCTATGTGAGCGGACCTACACGATGCGGGACCTCCATCGACCTCCAG  2133

358  SerGluValAspProArgPheMetGluArgGluHisIlePheArgGluArgGluAs  374        657  SerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuPr  674
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1234 TCAGAGGTAGACCCGGTTCATGGAGCGGGAACACCTCATATTCCGCGAGA  1283      2134  TCCTTCATGGCTCGAGGCCTTCAACAGCTCTATGGACATGGCTCGCCTCCCC  2183

374  nThrLeuAlaThrLysSerAlaIleGluGluThrTyrMetArgLeutIleGlyGlnL  391        674  oSerProThrLysProGlyProProProProProGlyGlyLysA  691
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1284 CACGCTCGCCACTAAAGCCATAGAAGACTATATGAGACTGATTGGCCAGA  1333      2184  CTCCCCAACCCAGGAGAAACCCCCGGCCCCCCCCGTGGGGTAAAG  2233

391  ysTyrLeuLysAspAlaIleGlyLysGluPheIleArgAlaLeuTyrGluSer  407        691  spLeuPheTyrValSerArgProLeuAlaArgSerProAlaTyr  707
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1334 AATACCTTCAAGGATGCCATTGGGAGTTCATCCGGGCTCTGTATGAATCT  1383      2234  ACCTGTTCTATGTGAGCCGGGCCACCACTGGCCCGGTCCTCCCCAGCATAC  2283

408  GluGluAsnCysGluValAspAspIleLysCysThrAlaSerLeuAl  424        708  CysThrSerSerSerAspIleThrGluProGluGlnLeuMetLeuSerVa  724
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1384 GAGGAGAACTGTGAAGTAGACCCCATCAAGTGCACAGCGTCCAGTCTGGC  1433      2284  TGCACCAGCAGCTCCGACATCACAGAGCCGGAGCAGATGCTGAGTGT  2333

424  aGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysV  441        724  lAsnLysSerValSerMetLeuAspLeuGlnGlnGlyLyApGlyProGlyGlyA  741
     ||||||||||||||||||||||||||||||||||||||||||||||||||                    |||||||||||||||||||||||||||||||||||||||||||||||||
1434 AGAGCACCAGGCCAACCTGCGGATGTGCTGTGAGTTGGCCCTGTGCAAGG  1483      2334  CAACAAGAGTGTGTCCATGCTGGACCTGCAGCAGGGCACGGGGCCTGGGGGGCC  2383

                                                                            741  rgLeuAsnSerSerSerValSerArgAsnLeuSerAsnLeuAlaLaValGlyAspLeuLeu  757
```

```
2384 |||||||||||||||||||||||||||||||||||
     GCCTAACACAGTAGTGTTCAACCTGGCAGCTGTTGGGGACCTGTTG 2433

 758 HisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaPr 774
     |||||||||||||||||||||||||||||||||||||||||||||||||
2434 CACTCAAGCCAGGCTTCACTGACAGCAGCCCTTGGGGTGCGGCTGCACC 2483

 774 oAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAlaGlyM 791
     |||||||||||||||||||||||||||||||||||||||||||||||||
2484 TGCCGGCGCGCCTCTCCAAGGGAGTGGCTCTTCCATCACAGCCGCCGCA 2533

 791 etArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnGln 807
     |||||||||||||||||||||||||||||||||||||||||||||||||
2534 TGCGCCCTCAGCCAGATGGGTGTCACTACGGATGGTGTCCCCGCCCAGCAA 2583

 808 LeuArgIleProLeuSerProPheGlnAsnProLeuPheHisMetAlaAlaAs 824
     |||||||||||||||||||||||||||||||||||||||||||||||||
2584 CTGCGCCATCCCTCTTTCCTTCCAGAACCCTCTTCCATATGGCTGCCGA 2633

 824 pGlyProGlyProAlaGlyHisGlyHisSerSerGlyHisSerGlyProP 841
     |||||||||||||||||||||||||||||||||||||||||||||||||
2634 TGGACCAGGGCCCCCAGCAGGCCATGGAGGAGCAGTGGCCATGGTCCAC 2683

 841 roSerHisHisHisHisHisHisHisHisHisHisArgGlyGlyGlu 857
     |||||||||||||||||||||||||||||||||||||||||||||||||
2684 CTTCCTCCCATCACCACCACCACCACCATCACCACCGAGGGGGAGAA 2733

 858 ProProGlyAspThrPheAlaProPheHisGlyTyrSerLysSerGluAs 874
     |||||||||||||||||||||||||||||||||||||||||||||||||
2734 CCCCCAGGGAGACACTTTTGCCCCGTTCCATGGCTATAGCAAGAGCGAGGA 2783

 874 pLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHisSerH 891
     |||||||||||||||||||||||||||||||||||||||||||||||||
2784 CCTCTCTACAGGGGTCCCTAAGCCCCTAGCCGCCTCCATCCTTCACAGCC 2833

 891 isSerTyrSerAspGluPheGlyProSerGlyThrAspPheThrArgArg 907
     |||||||||||||||||||||||||||||||||||||||||||||||||
2834 ACAGCTACAGTGATGAGTTTGGACCCTCTGGTACTGATTTTACCCGTCGG 2883

 908 GlnLeuSerLeuGlnAsnLeuMetLeuSerProProGlnIl 924
     |||||||||||||||||||||||||||||||||||||||||||||||||
2884 CAGCTCTCACTTCAGGACAACCTACAGCACATCCTCTCCCCGCCCCAGAT 2933

 924 eThrIleGlyProGlnArgProAlaProSerGlyPProGlyGlyGlySerG 941
     |||||||||||||||||||||||||||||||||||||||||||||||||
2934 CACCATCGGTCCCCAGAGGCCAGCTCCTCAGGGCAGGAGGGGGCAGTG 2983

 941 lyGlyGlyGlyGlyGlyGlyGlnProProLeuGlnArg 957
     |||||||||||||||||||||||||||||||||||||||||||||||||
2984 GTGGGGCAGTGGTGGGGCGGTTGGGGCGGTCAGCCACCTCCCTTGCCAGAGG 3033

 958 GlyLysSerGlnLeuThrValSerArgAlaAlaGlnLysProArgProSe 974
     |||||||||||||||||||||||||||||||||||||||||||||||||
3034 GGCAAATCTCAGCAGTTGACAGTGAGTGCTGCCCAGAAACCCGGCCGTC 3083

 974 rSerGlyAsnLeuGlnLnSerProGluProSerTyrGlyProAlaArgP 991
     |||||||||||||||||||||||||||||||||||||||||||||||||
3084 CAGCGGGAACCTATTGCAGTCCCCGGAACCAAGTTATGGTCCTGCCCGTC 3133

 991 roArgGlnGlnSerLeuSerLysGluGlyGlyGlyGly 1007
     |||||||||||||||||||||||||||||||||||||||||||||||||
3134 CACGGCCAACAGAGCCTCAGCAAAGAGGGCAGCATTGGGGGCAGCGGGGGC 3183

1008 SerGlyGlyGlyGlyGlyGlyLysLeuLysProSerIleThrLysSerGlnHi 1024
     |||||||||||||||||||||||||||||||||||||||||||||||||
3184 AGCGGTGGCGGAGGGGGTGGGGGCGGCTCAAGCCCTCCATCACCAAGCAGCA 3233

1024 sSerGlnThrProSerProHisLeuAsnProThrMetProAlaSerGluArgT 1041
     |||||||||||||||||||||||||||||||||||||||||||||||||
3234 TTCCCAGACTCCATCCACGCTGAACCCCACGATGCCGGCCTCGGAGCGGA 3283

2334 orValAlaTrpValSerAsnMetProHisLeuSerLeuSerAlaAspIleGluSer 1057
     |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
3284 CTGTAGCCTGGGTGTCCAATATGCCTCACCTGTCCGCTGACATCGAGAGT 3333

1058 AlaHisIleGluArgGluGluThrGluTyrLysLeuGluTyrSerLysSerMe 1074
     |||||||||||||||||||||||||||||||||||||||||||||||||
3334 GCACACATTGAGCGGGAAGAGTACAAGCTGGAGGAGTACTCGAAGTCCAT 3383

1074 tAspGluSerArgLeuAspAspArgValLysGluTyrGluGluIleHisiS 1091
     |||||||||||||||||||||||||||||||||||||||||||||||||
3384 GGACGAGAGCCGACTGGACAGGGTGAAGGAGTACGAGGAGATCCACT 3433

1091 erLeuLysGluArgLeuHisMetSerAsnArgLysLeuGluGluTyrGlu 1107
     |||||||||||||||||||||||||||||||||||||||||||||||||
3434 CACTGAAGGAGAAGGCTACACATGTCCAACCGAAGCTGGAAGAGTACGAG 3483

1108 ArgArgLeuSerGlnGluGlnLnThrSerLysIleLeuMetGlnTy 1124
     |||||||||||||||||||||||||||||||||||||||||||||||||
3484 CGGAGGCTGCTGTCCCAGGAAGAGCAGACCAAGATCCTGATGCAGTA 3533

1124 rGlnAlaArgLeuGlnSerGluGluLysArgLeuArgGlnGlnGlnnValG 1141
     |||||||||||||||||||||||||||||||||||||||||||||||||
3534 CCAAGCCCGCCTTGGAGCGAGAGGAGAAGCGCTTGAGGCAGCAGCAGGTGG 3583

1141 luLysAspSerGlnIleLeuLysSerIleIleGlyArgLeuMetLeuValGlu 1157
     |||||||||||||||||||||||||||||||||||||||||||||||||
3584 AGAAGGACTCCCAGATCAAGAGCATCATTGGCAGGCTGATGCTGGTGGAG 3633

1158 GluGluLeuArgArgAspHisProAlaMetAlaGluProLeuProGluPr 1174
     |||||||||||||||||||||||||||||||||||||||||||||||||
3634 GAGGAGCTGCGCCGGGACCACCCCATGGCTGAGCCGCTGCCTGAACC 3683

1174 oLysLysArgLeuLeuAsnAlaGlnAlaArgGlySerPheProProTrpValG 1191
     |||||||||||||||||||||||||||||||||||||||||||||||||
3684 CAAGAAGAGGCTGCTGAACGCTCAAGAGAGCAGCTTCCCCCCTTGGGTCC 3733

1191 lnGlnThrArgVal...ArgTrpProHisLeuGlyThrAlaTrpProPro 1206
     |||||||||||||||||||||||||||||||||||||||||||||||||
3734 AACAAACCCGGTGACGCTGGCCCCACCTTGGAACGGCCTGGCCCCCC 3783

1207 GlnProHisProProHisProProGlyCysArgSerGlnArgThrAlaSerSe 1223
     |||||||||||||||||||||||||||||||||||||||||||||||||
3784 CAGCCCCCACCCCCACCCCGGCCTGCAGATCACAAGAGAACGGCGAGTC 3833

1223 rGlyThrProGlnThrThrSerProProSerIleThrAspLeuLeuProC 1240
     |||||||||||||||||||||||||||||||||||||||||||||||||
3834 CGGAACACCGCAGACCACTAGCCCACCCAGCATCACAGACCTCCTTCCCT 3883

1240 ysAlaProTyrProGlyTyrProProSerValThrAspLeuLeuProSerAla 1256
     |||||||||||||||||||||||||||||||||||||||||||||||||
3884 GTGCACCCTACCCCGGCCACCGGTGTTGGGGCCAGCCACGTCACCAAGGGAC 3933

1257 ProAspProGlyThrSerProThrThrArgThrGlyArgHisGlnGlyTh 1273
     |||||||||||||||||||||||||||||||||||||||||||||||||
3934 CCCGACCCTGGAACATCACCACCAACCACCAGGACTGACGTCACCAAGGGAC 3983

1273 rAlaGlyLeuSerProLeuThrProProTrpGlyThrHisLeuSerThrP 1290
     |||||||||||||||||||||||||||||||||||||||||||||||||
3984 AGCGGGATTGTCCCTTAACGCCCTCCTTGGGCACCCATCTGTCAACCC 4033

1290 roLeuLeuHisSerArgArgGluSerGlyThrLeuSerGlyCysProLeuThr 1306
     |||||||||||||||||||||||||||||||||||||||||||||||||
4034 CACTTGCTCCATTCCAGGAGGGAGAGTGGGACCCTCAGCTGCCCTCTCACC 4083

1307 ProGlyHisHisLeuProHisThrAspProPheThrLeuGlyCysTyrPr 1323
     |||||||||||||||||||||||||||||||||||||||||||||||||
4084 CCAGGACACCTACCCCCACCACAGCCCCTTCACTCTGGGGTGCTATCC 4133

1323 oHisPro 1325
     ||||||
4134 CCATCCT 4140

seq_name: gb_ro:AF058789

seq_documentation_block:
```

```
LOCUS        AF058789      4539 bp    mRNA            ROD        14-SEP-2000
DEFINITION   Rattus norvegicus SynGAP-a mRNA, complete cds.
ACCESSION    AF058789
VERSION      AF058789.2  GI:10122137
KEYWORDS     .
SOURCE       Norway rat.
ORGANISM     Rattus norvegicus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
             Rattus.
REFERENCE    1  (bases 1 to 4539)
AUTHORS      Kim,J.H., Liao,D., Lau,L.F. and Huganir,R.L.
TITLE        SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
             protein family
JOURNAL      Neuron 20 (4), 683-691 (1998)
MEDLINE      98240917
REFERENCE    2  (bases 1 to 4539)
AUTHORS      Kim,J.H. and Huganir,R.L.
TITLE        Direct Submission
JOURNAL      Submitted (09-APR-1998) Neuroscience, Johns Hopkins Medical
             Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
COMMENT      On Sep 14, 2000 this sequence version replaced gi:3065888.
FEATURES             Location/Qualifiers
     source          1..4539
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="hippocampus"
     CDS             1..3855
                     /note="neuronal RasGAP; contains PH domain, C2 domain, GAP
                     domain in the N-terminal region, and a T/SXV motif at the
                     C-terminus"
                     /codon_start=1
                     /product="SynGAP-a"
                     /protein_id="AAC63510.2"
                     /db_xref="GI:10122138"
                     /translation="MSYAPFRDVRGPPMHRTQYVHSPYDRPGWNPRFCIISGNQLLML
                     DEDEIHPLLIRDRRSESSRNKLLRRTVSVPVEGRPHGEHEYHLGRSRRKSVPGGKQYS
                     MEAAPAAPFRPSQGFLSRRLKSSIKRTKSQPKLDRTSSFRQILPRFRSADHDRARLMQ
                     SFKESHSHESLLSPSSAAEALELNLDEDSIIKPVHSSILGQERPFEVTSSGTKCFAC
                     RSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELPKKRYYCELCLDDMLY
                     ARTTSKPRSASGDTVFWGEHFEFNNLPAVRALRLHLYRDSDKKKKDKRAGYVGLVVVP
                     VATLAGRHFTEQWYPVTLPTGSGGSGMGSGGGSGGGSGGGSGKGKGCPAVRLKARYQ
                     TMSILPEMELYKEFAEYVTNHYRMLCAVLEPALNVKGKEEVASALVHILQST GKADFL
                     SDMAMSEVDRRMEREHLIFRENTLATKAIEEYMRLIGQKYLKDAIGEFIRALYESEEN
                     CEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKEVFASWRLRCAERGR
                     EDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTLTLIAKVIQNLANFSKF
                     TSKEDPLGFMNEFLELEWGSMQQFLYEISNLDTLFNSSFPEGYIDLGRELSTLHALLW
                     EVLPQLSKEALLKLGPLPRLLSDISTALRNFNIQRQPSRQSERARSQPMVLRGPSAEM
                     QGYMMRDLNSIDLQSFMARGLNSSMDMARLPSPTKEKPPPPPGGGKDLFYVSRPPL
                     ARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGDGPGGRLNSSSVSNLAAVGDLLH
                     SSQASLTAALGLRAPAGRLSQGSGSSITAAGMRLSOMGVTTDGVPAQLRIPLSFQN
                     PLFHMAADGPGPPAGHGGSSGHGPPSSHHHHHHHHHRGGEPPGDTFAPFHGYSKSED
                     LSTGVPKPPAASILHSHSYSDEFGPSGTDFTRRQLSLQDNLQHMLSPPQITIGPQRPA
                     PSGPGGSGGSGGGGQPPLQRGKSQQLTVSAAQKPRPSSGNLLQSPEPSYGPAR
                     PRQQSLSKEGSIGGSGGSGGGGGLKPSITKQHSQTPSTLNPTMPASERTVAWVSNM
                     PHLSADIESAHIEREEYKLKEYSKSMDESRLDRVKEYEEIHSLKERLHMSNRKLEEY
                     ERRLLSQEEQYSKILMQYQARLEQSEKRLRQQQVERDSQIKSIIGRLMLVEEELRRDH
                     PAMAEPLPEPKKRLLDAQLLIR"
BASE COUNT     992 a   1448 c   1290 g    809 t
ORIGIN
```

```
alignment_scores:
     Quality:    6858.00          Length:     1346
     Ratio:      5.219            Gaps:          2
Percent Similarity: 97.623    Percent Identity: 97.548

alignment_block:
US-09-294-298-6 x AF058789

Align seg 1/1  to: AF058789  from: 1  to: 4539

     8  AlaProHisProPheAlaProGlnGlnGlyPheLeuSerArgArgLe  24
```

```
          |||||||                  ||::::|||||||||||||
316  GCCCCGCTGGCGCCCCCCCTTCCGGCCCCTCGCAGGCTTCCTGAGCCGGAGGCT  365

24  uLysSerSerIleLysArgThrLysSerGlnProLysLeuAspArgThrS  41
         ||||||||||||||||||||||||||||||||||||||||||||||||
366  AAAAAGCTCCATCAAACTACAAAGTCACAACCCAAACTTGACCGGACCA  415

41  erSerPheArgGlnIleLeuProArgPheArgSerAlaAspHisAspArg  57
         ||||||||||||||||||||||||||||||||||||||||||||||||
416  GCAGCTTTCGACAGATCCTGCCTCCGCTTCCGAAGTGCTGACCATGACCGG  465

58  AlaArgLeuMetGlnSerPheLysGluSerHisSerHisGluSerLeuLe  74
         ||||||||||||||||||||||||||||||||||||||||||||||||
466  GCCCGGCTGATGCAGAGCTTCAAGGAGTCTCACTCCCATGAGTCCCTGCT  515

74  uSerProSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAspS  91
         ||||||||||||||||||||||||||||||||||||||||||||||||
516  GAGTCCCAGCAGTGCTGTCGAGGCCCTGGAGCTCAACCTGGATGAAGACT  565

91  erIleIleLysProValHisSerSerIleLeuGlyGlnGluPheCysPhe  107
         ||||||||||||||||||||||||||||||||||||||||||||||||
566  CCATTATCAAGCCAGTACACAGCTCCATCCTGGGCCAGGAGTTCTGCTTT  615

108  GluValThrThrSerSerGlyThrLysCysPheAlaCysArgSerAlaAl  124
         ||||||||||||||||||||||||||||||||||||||||||||||||
616  GAGGTAACAACATCGTCTGGGACAAAATGTTTGCCTGTCGGTCTGCAGC  665

124  aGluArgAspLysTrpIleGluAsnLeuGlnArgAlaValLysProAsnL  141
         ||||||||||||||||||||||||||||||||||||||||||||||||
666  CGAAAGGGACAAATGGATTGAGAATCTACAGAGGCTGTGAAACCCAACA  715

141  ysAspAsnSerArgArgValAspAsnValLeuLysLeuTrpIleIleGlu  157
         ||||||||||||||||||||||||||||||||||||||||||||||||
716  AGGACAACAGCCGCCGGTAGATAACGTGCTGAAACTATGGATCATAGAA  765

158  AlaArgGluLeuProProLysLysArgTyrTyrCysGluLeuCysLeuAs  174
         ||||||||||||||||||||||||||||||||||||||||||||||||
765  GCTCGAGAGCTGCCCCCAAGAAGCGATATTACTCGAGTTATGCCTGGA  815

174  pAspMetLeuTyrAlaArgThrThrSerLysProArgSerAlaSerGlyA  191
         ||||||||||||||||||||||||||||||||||||||||||||||||
816  CGACGTCTATGCACGGACCACCTTCCAAGCCCCGCTCAGCCTCAGGAG  865

191  spThrValPheTrpGlyGluHisPheGluPheAsnAsnLeuProAlaVal  207
         ||||||||||||||||||||||||||||||||||||||||||||||||
866  ACACTGTCTTTTGGGGCGAGCACTTCGAGTTAACAACCTGCCTGCGTC  915

208  ArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysLysArgLySLy  224
         ||||||||||||||||||||||||||||||||||||||||||||||||
916  CGGGCGCTGCGGCTGCATCTGACCTGACCGGACAAAAAGCGGAAGAA  965

224  sAspLysAlaGlyTyrValGlyLeuValThrValProValAlaThrLeuA  241
         ||||||||||||||||||||||||||||||||||||||||||||||||
966  GGACAAGGCAGGCTACGTGGCCTGGCCTGGCTGTTCCAGTGACCACCCTGG  1015

241  laGlyArgHisPheThrGluGlnTrpTyrProValThrLeuProThrGly  257
         ||||||||||||||||||||||||||||||||||||||||||||||||
1016  CTGGGCGCCACCTTCACAGAGCAGTGTACCCCGTGACCCTGCCAACAGGA  1065

258  SerGlyGlySerGlyGlyMetGlySerGlyGlyGlyGlySerGlyGlyGl  274
         ||||||||||||||||||||||||||||||||||||||||||||||||
1066  AGTGGGGGCTCTGGGGGGTATGGGCTCGGGGGCAGGGGGGTCAGGGGG  1115

274  yGlySerGlyGlyLysGlyGlyLysGlyCysProAlaValArgLeuLysA  291
         ||||||||||||||||||||||||||||||||||||||||||||||||
1116  CGGCTCAGGGGGCAAAGGGAAAGGAGGCTGTCCTGCTGGCTGAAGG  1165

291  laArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLysGluPhe  307
         ||||||||||||||||||||||||||||||||||||||||||||||||
1166  CCCGTTACCAGCACAATGAGTATCCTGCCCATGGAGCTATATAAGGAGTTT  1215

308  AlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaValLeuGluPr  324
         ||||||||||||||||||||||||||||||||||||||||||||||||
1216
```

Fri Jan 19 06:20:05 2001

(Two-column nucleotide/protein sequence alignment listing)

Right column:

```
624  nSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSerAlaG    641
2166 GAGCGAGGCGCGCTCGGGTTCAGCCCATGGTGCTGCGCGGCCGTCAGCCG   2215

641  luMetGlnGlyTyrMetArgAspLeuAsnSerSerIleAspLeuGln       657
2216 AGATGCAGGGCTACATGATGCGGGACCTCAACAGCTCCATCGACCTTCAG   2265

658  SerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuPr    674
2266 TCCTTCATGGCTCGAGGCCTCAACAGCTCTATGGACATGGCTCGCCTCCC   2315

674  oSerProThrLysGluLysProProProProProGlyGlyGlyLysA       691
2316 CTCCCCAACCAAGGAGAAACCCCGCGCCCCCCTCCCGGTGGGGGTAAAG    2365

691  spLeuPheTyrValSerArgProProLeuAlaArgSerProAlaTyr       707
2366 ACCTGTTCTATGTGAGCCGGCCACCACTGGCCCGGTCCTCCCAGCATAC    2415

708  CysThrSerSerSerAspIleThrGluProGlnLysMetLeuSerVa       724
2416 TGCACGGAGCAGCTCGGACATCACAGAGCCGGAGCAGAAGATGCTGAGTGT  2465

724  lAsnLysSerValSerMetLeuAspLeuGlnGlyGlyAspGlyProGlyGlyA 741
2466 CAACAAGAGTGTGTCCATGCTGGACCTGCAGGGCGACGGGCCTGGGGGCC  2515

741  rgLeuAsnSerSerSerValSerAsnLeuAlaAlaAlaValGlyAspLeuLeu 757
2516 GCCTTAACAGCAGTAGTGTTTCCAACCTGGCAGCTGCTGTTGGGGACCTGTTG 2565

758  HisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaPr    774
2566 CACTCAAGCCAGGCTTCACTGACAGCCGCCTTGGGTTGCGGCCTGCACC    2615

774  oAlaGlyArgLeuArgGlnGlySerGlySerSerIleThrAlaAlaGlyM    791
2616 TGCCGGGCCGCCTCTCCCAAGGAGTGCTCTTCCATCACAGCAGCCGGCA    2665

791  etArgLeuSerGlnGlnMetGlyValThrThrAspGlyValProAlaGlnGln 807
2666 TGCGGCCTCAGCCAGATGGGTGTCACTACGGATGGTGTCCCCGCCCAGCAA 2715

808  LeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaAs    824
2716 CTGCGCATCCCTCTTCCTTCCAGAACCCTCTTCTTCCATATGGCTGCCGA  2765

824  pGlyProGlyProProAlaGlyHisGlyLeuSerGlyHisGlyProP       841
2766 TGGACCAGGGCCCCCAGCAGGCCATGAGGGAGCAGTGGCCATGGTCCAC   2815

841  roSerSerHisHisHisHisHisHisArgGlyGlyGlu               857
2816 CTTCCTCCCATCACCACCATCACCAGGGGAGGAGAA                 2865

858  ProProGlyAspThrPheAlaProHisGlyTyrSerLysSerGluAs       874
2866 CCCCCAGGGGACACTTTTGCCCCGTTCCATGCCTATAGCAAGAGCGAGGA  2915

874  pLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHisSerH    891
2916 CCTCTCTACAGGGGTCCCTAAGCCCCCTGCGGCCTCCATCCTTCACAGCC  2965

891  isSerThrTyrSerAspGluPheGlyProSerGlyThrArgArgArg       907
2966 ACAGCTACAGTGATGAGTTTGGACCCTCTGGTACTGATTTTACCCGTCGG  3015

908  GlnLeuSerLeuGlnAsnAsnLeuGlnHisMetLeuSerProGlnIl      924
3016 CAGCTCTCACTTCAGGACAACCTACAGCACACAATGCTCTCCCCGCCCAGAT 3065
```

Left column:

```
1216 GCAGAATATGTGACCAACCACTACCGCATGCGTGTGCGCTGCTGGAGCC   1265

324  oAlaLeuAsnValLysGlyLysGluValAlaSerAlaLeuValHisI      341
1266 CGCCCTCAAATGTCAAGAGCCAAGGAGGTGCTAGTGCCACTGGTTCACA   1315

341  leLeuGlnSerThrGlyLysAlaLysAspPheLeuSerAspMetAlaMet   357
1316 TCCTGCAAAGCACAGGCAAGGCCAAGGACTTCCTTTCAGACATGGCCATG  1365

358  SerGluValAspArgPheMetGluArgGluHisLeuIlePheArgGluAs   374
1366 TCAGAGGTAGACCGGTTCATGGAGCGGGAACACCTCATATTCCGCGAGAA  1415

374  nThrLeuAlaThrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLnL 391
1416 CACGCTCGCCACTAAAGCCATAGAAGAGTATATGAGACTGATTGGCCAGA  1465

391  ysTyrLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyrGluSer   407
1466 AATACCTCAAGGATGCCATTGGGGAGTTCATCCGGGCTCTGTATGAATCT  1515

408  GluGluAsnCysGluValAspProIleLeuCysSerThrAlaSerLeuAl   424
1516 GAGGAGAACTGTGAAGTAGACCCCATCAAGTGCAGCACAGCCTCCAGTCTGGC 1565

424  aGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysSV  441
1566 AGAGCACCAGGCCAACCTGCGGATGTGCTGTGAGTTGGCCCTGTGCAAGG  1615

441  alValAsnSerHisCysValPheProArgGluLeuLysGluValPheAla   457
1616 TGGTCAACTCCCATTGCGTGTTCCCGAGGGAGCTGAAGGAGGTCTTTGCA  1665

458  SerTrpArgLeuArgCysAlaGluArgGlyArgGluAlaAspPAr        474
1666 TCATGGCGGCTGCGCTGCGCAGAGCGCGGGCGGGAGGCAATTGCTGACAG  1715

474  gLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetS   491
1716 GCTGATCAGCGCCTCCCTCTTCCTGCGCTTCCTGTGCCCCGGCCATCATGT 1765

491  erProSerLeuPheGlyLeuMetGlnGluTyrProAspGlnLnThrSer    507
1766 CGCCCAGTCTGTTTGGACTGATGCAGGAGTACCCAGATCAGCAGCCTTCA  1815

508  ArgThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSe   524
1816 CGAACCCTCACCCTCATCGCCAAGGTTATCCAGAACCTGGCCAACTTTTC  1865

524  rLysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuG   541
1866 CAAGTTTACCTCAAAGGAGGACTTTCTGGGCTTCATGAACGAGTTTCTGG 1915

541  luLeuGluTrpGlySerMetGlnGluPheLeuTyrGlnIleSerAsnLeu   557
1916 AGCTGGAGTGGGGTTCTATGCAGGAATTCTTGTATCAGATATCCAACCTG  1965

558  AspThrThrAsnSerSerPheSerGluGlyTyrIleAspLeuGlyTyrAr   574
1966 GACACCACTAACAGCAGCAGTTTTGAGGGCTACATAGACTTGGGCCG     2015

574  gGluLeuSerThrLeuHisAlaLeuLeuTrpGluValProGlnIleLeus   591
2016 CGAGCTCTCCACACTTCACGCCCTGCTCTGGGAGGTGCTCCCCCAGCTCA  2065

591  erLysGluAlaLeuLeuLysLeuGlyProLeuProLeuLeuSerAsp      607
2066 GCAAGGAAGCCCTCCTGAAGCTGGGCCCGCTGCCCCTGCTCTCAGCGAC   2115

608  IleSerThrAlaLeuArgAsnProAsnIleGluArgGlnSerProSerArgGl 624
2116 ATCAGCACAGCCCTGAGGAACCCTAACATCCAAAGGCAGCCGAGCCGCCA  2165
```

Backfromlate Seg10N6. (handwritten) — 6

```
 924 eThrIleGlyProGlnArgProAlaProSerGlyProGlyGlyGlySerG  941
     ||||||||||||||||||||||||||||||||||||||||||||||||
3066 CACCATCGGTCCCCAGAGGGCCAGCTCCCCAGGGCCCGGAGGGGGCAGTG 3115

 941 lyGlyGlySerGlyGlyGlyGlyGlyProProLeuGlnArg           957
     |||||||||||||||||||||||||||||||||
3116 GTGGGGCAGTGTGGTGGGGGCGGTGGGGGCCACCTCCCTTGCCAGAGG   3165

 958 GlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysProArgProSe  974
     ||||||||||||||||||||||||||||||||||||||
3166 GGCAAATCTCAGCAGTTGACAGTGAGTGCTGCCCAGAAACCCCGGCCGTC 3215

 974 rSerGlyAsnLeuLeuGlnSerProGluSerTyrGlyProAlaArgP     991
     |||||||||||||||||||||||||||||||||
3216 CAGCGGGAACCTATTGCAGTCCCCGGAACCAAGTTATGGTCCTGCCCGTC 3265

 991 roArgGlnGlnSerLeuSerLeuGlyGluSerGlyGlySerGlyGly    1007
     ||||||||||||||||||||||||||||||||||
3266 CACGCAACAGAGCCTCAGCAAAGAGGGCAGCATTGGGGGCAGCGGGGGC  3315

1008 SerGlyGlyGlyGlyGlyLeuLeuLysProSerIleThrLysGlnHi    1024
     ||||||||||||||||||||||||||||||||
3316 AGCGGTGGCGGCGGAGGGGGTGGGGGGCTCAAGCCCTCCATCACCAAGCACA 3365

1024 sSerGlnThrProSerThrLeuAsnProThrMetProAlaSerGluArgT 1041
     ||||||||||||||||||||||||||||||||
3366 TTCCCAGACTCCATCCACCGGTGAACCCCACGATGCCGGCCTCCGAGCGGA 3415

1041 hrValAlaTrpValSerAsnMetProHisLeuSerAlaAspIleGluSer 1057
     ||||||||||||||||||||||||||||||||
3416 CTGTAGCCTGGGTGTCCAATATGCCTCACCTGTCCGCTGACATCGAGAGT 3465

1058 AlaHisIleGluArgGluGluLysLysLeuLysTyrLysGluLysSerMe 1074
     ||||||||||||||||||||||||||||||||
3466 GCACACATTGAGCGGGAAGAGTACAAGCTGAAGGAGTACTCGAAGTCCAT 3515

1074 tAspGluSerArgLeuAspArgValLysSerGluTyrGluGluIleHisS 1091
     ||||||||||||||||||||||||||||||||
3516 GGACGAGAGCCGACTGGACAGGGTGAAGAGTACGAGGAGGAGATCCACT  3565

1091 erLeuLysGluArgLeuHisMetSerAsnArgLysLeuGluThrTyrGlu 1107
     ||||||||||||||||||||||||||||||||
3566 CACTCGAAGGAAAGGCTACACATGTCCAACCGGAAGCTGGAAGAGTACCAG 3615

1108 ArgArgLeuLeuSerGlnGluGluGlnThrSerIleIleLeuMetGlnTy 1124
     ||||||||||||||||||||||||||||||||
3616 CGGAGGCTGCTGTCTCAGGAAGAGCAGACCAGCAAGATCCTGATGCAGTA 3665

1124 rGlnAlaArgLeuGlnGlnSerGluLysArgLeuArgGlnGlnValGlnG 1141
     ||||||||||||||||||||||||||||||||
3666 GCAGGCTCGGCTGCAGCAGTCTGAGAAGCGGCTGCGGCAGCAGGTGCAGG 3715

1141 luLysAspSerGlnIleLeuLysSerIleIleGlyArgLeuMetLeuValGlu 1157
     ||||||||||||||||||||||||||||||||
3716 AGAAGGACTCTCAGATCAAGAGCATCATTGGCAGGCTGATGCTGGTGGAG 3765

1158 GluGluLeuArgArgAspHisProAlaMetAlaGluProLeuProGluPr 1174
     ||||||||||||||||||||||||||||||||
3766 GAGGAGCTGCGCCGGGACCACCCCGCCATGGCTGAGCCCCTGCCTGAACC 3815

1174 oLysLysAspLeuAspAlaGln............                 1182
     |||||||||||||||||||||
3816 CAAGAAGGAGCTGCTCGACGCTCAGCTCCTCATCAGGTAATTCTCCTGGT 3865

1182 ..............                                     1182
3866 TCCGCTTTGACCACCGGCCGGAGGACACTGGGGGAGGTGACTCCGGACCAC 3915

1183 .....ArgGlySerPheProProTrpValGlnGlnThrArgVal...Ar 1196
     ||||||||||||||||||    ||
3916 TGCAGGAGGCAGCTTCCCCCCTTGGGTCCAACAAACCCGCGTGTGACG   3965

1196 gTrpProHisLeuGlyThrAlaTrpProProGlnProGlnProHisP   1213
```

```
3966 CTGGCCCACCTTGGAACGGCCTGCCCCCCAGCCCACCCCCACC        4015
     ||||||||||||||||||||||||||||||||||||||
1213 roGlyCysArgSerGlnArgThrAlaSerSerGlyThrProGlnThrThr 1229
4016 CCGGCTGCAGATCACAGAGAACGGCAGTTCCGAGAACACGCAGCACCACT 4065

1230 SerProProSerIleThrAspLeuProLeuLeuProCysAlaProTyrProGlyPr 1246
     ||||||||||||||||||||||||||||||||
4066 AGCCCACCCAGCATCACAGACCTCCTTCCCTGTGCACCCTACCCCGGCCC 4115

1246 oProSerValThrArgAspLeuProSerAlaProAspProGlyThrSerP 1263
     ||||||||||||||||||||||||||||||||
4116 ACCCAGCGTCACAGACCCTCCAGTGCACCCGACCCTGGAACATCAC     4165

1263 roThrThrArgThrGlyArgHisGlnGlyThrHisAlaGlyLeuSerProLeu 1279
     ||||||||||||||||||||||||||||||||
4166 CAACCACCAGGACTGGACGTCACCAAGGACAGCAGCGGGATTGTCTCCCTTA 4215

1280 ThrProProTrpGlyThrHisLeuSerThrProLeuLeuHisSerArgAr 1296
     |||||||||||||||||||||||    |||||||
4216 ACGCCTCCCTTGGGGCACGCT.CTGTCAACCCCACTGCTCCATTCCAGGAG 4264

1296 gGluSerGlyThrLeuSerProLeuThrProGlyHisHisLeuProH   1313
     ||||||||||||||||||||||||||||||||
4265 GGAGAGTGGGACCCTCCAGCTGCCCCTCCACCCAGGACACCACCTACCCC 4314

1313 isThrAspProPheThrLeuGlyCysTyrProHisPro             1325
     |||||||||||||||||||||||||
4315 ACACAGACCCCCTTCACTCTGGGGTGCTATCCCACATCCT           4352
```

seq_name: gb_ro:AF048976

seq_documentation_block:
```
LOCUS       AF048976     4063 bp    mRNA            ROD       27-MAR-1998
DEFINITION  Rattus norvegicus synaptic ras GTPase-activating protein p135
            SynGAP mRNA, complete cds.
ACCESSION   AF048976
VERSION     AF048976.1  GI:2935447
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4063)
  AUTHORS   Chen,H.-J. and Kennedy,M.B.
  TITLE     Identification and cloning of a novel 130 kd protein containing a
            ras GTPase-activating domain from the rat forebrain postsynaptic
            density
  JOURNAL   (in) SOC. NEUROSCI. ABSTR.: 1466;(1997)
REFERENCE   2  (bases 1 to 4063)
  AUTHORS   Chen,H.-J., Rojas-Soto,M. and Kennedy,M.B.
  TITLE     A synaptic Ras GTPase-activating protein (p135 SynGAP) inhibited by
            CaM kinase II
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 4063)
  AUTHORS   Chen,H.-J. and Kennedy,M.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-1998) Division of Biology, California Institute
            of Technology, 1200 E. California Blvd., MC 216-76, Pasadena, CA
            91125, USA
FEATURES             Location/Qualifiers
     source          1..4063
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     CDS             122..4003
                     /function="prominent substrate for endogenous CaMKII; can
                     activate intrinsic ras GTPase activity"
                     /note="synaptic ras-GAP; N-terminal encodes putative PH
                     domain, C2 domain, and ras-GAP domain; C-terminal encodes
                     proline-rich region, stretch of 10 histidine residues, and
```

```
t/SXV motif; enriched in the forebrain postsynaptic
density fraction"
/codon_start=1
/product="synaptic ras GTPase-activating protein p135
SynGAP"
/protein_id="AAC08071.1"
/db_xref="GI:2935448"
/translation="MSYAPFRDVRGPPMHRTQYVHSPYDRPGWNPRFCIISGNQLLML
DEDEIHPLLIRDRRSESSRNKLLRRTVSVPVEGRPHGEHEYHLGRSRRKSVPGGKQYS
MEARAPAAPFRPSOGFLSRRLKSSIKRFKSQPKLDRTSSFRQILPRRSADHDRARLMQ
SFKESHSHSLLSLPSSAAEALELNLDEDSIIKPVHSSILGQEFCFEVTTSSGTKCFAC
RSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELPPKRRYYCELCLDDMLY
ARTTSKPRSASGDTVFWGEHFEFNNLPAVRALRLHLYRDSDKKRRKDKAGYVGLVTVP
VATLAGRHFTEQWYPVTLPTGSGSGGMGSGGGSGGSGKGKGCPAVRLKARYQ
TMSILPMELYKEFAEYVTNHYRMLCAVLEPALNVKGKEEVASALVHILQSTGKAKDFL
SDMAMSEVDRFMEREHLIFRENTLATRKAIEEYMRLIGQKYLKDAIGEFIRALYESEEN
CEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKEVFASWRLRCAEBRGR
EDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTLTLIAKVIQNLANFSKF
TSKEDFLGFMNEFLELEWGSMOQFLYEISNLDTLTNSSSFPGYIDLGRELSTLHALLW
EVLPQLSKEALLKLGPLPRLLSDISTALRNPNIQRQPSRQSERARSQPMVLRGPSAEM
QGIYMRDLNSSIDLQSFMARGLNSSMDMARLPSPTKEKPPPPPGGKDLFYVSRPPL
ARSSPAYCTSSSDITEPEQKMLSVNKVSMLDLQGDGPGGRLNSSSVSNLAAVDLLH
SSQASLTAALGLRPAPAGRLSQCSGSSITAAGMRLSQMGVTTDGVPAQQLRIPLSFQN
PLFHMAADGPGPPAGHGGSSGHGPPSSHHHHHHHHRGGEPPGDTFAPFHGYSKSED
LSTGVPKPPAASTLHSHYSDEFGPSGTDRRQLSLQDNLQHMLSPPQITIGPORPA
PSGPGGSGSGSGGGGSGQPPLQRGKSQQLTVSAAQKPRPSSGNLLQSPEPSYGPAR
PRQQSLSKEGSIGGSGGSGGGGGLKPSITKQHSQTPSTLNPTMPASERTVAWVSNM
PHLSADIESAHIEREEYKLKEYSKMDESRLDRVKEYEEIIHSLKERLHMSNRKLEEY
ERRLLSOEEOTSKILMOYQARLEQSEKRLRQQOVEKDSQIKSIIGRLMLVEEELRRDH
PAMAEPLPEPKKRLLDAQRGSFPPWVQQTRV"
```

BASE COUNT   880 a   1273 c   1177 g   733 t
ORIGIN

alignment_scores:
    Quality: 6292.50   Length: 1209
    Ratio: 5.222   Gaps: 1
Percent Similarity: 99.669   Percent Identity: 99.586

alignment_block:
US-09-294-298-6 x AF048976

Align seg 1/1 to: AF048976 from: 1 to: 4063

```
     8 AlaProHisPheProPheAlaProGlnGlnGlyPheSerArgArgIle      24
       |||||       |||||    ||::::|||||||||||||||||||||
   437 GCCCCCGCTGCGCCCTTCCGGCCGCTTCCTGAGCCGGAGGCT              486

    24 uLysSerIleLeuLysArgThrLysSerGlnProLysLeuAspArgThrS     41
       ||||||||||||||||||||||||||||||||||||||||||||||||
   487 AAAAAGCTCCATCAAACGTACAAAGTCACAACCCAAACTTGACCGGACCA     536

    41 erSerPheArgGlnIleLeuProArgPheArgSerAlaAspHisAspArg     57
       ||||||||||||||||||||||||||||||||||||||||||||||||
   537 GCAGCTTTCGACAGATCCTGCCTTCCGAGTGCTGACCATGACCGG          586

    58 AlaArgLeuMetGlnSerPheLysSerHisSerHisSerLeuLeu         74
       ||||||||||||||||||||||||||||||||||||||||||||||
   587 GCCCGGCTGATGCAGAGCTTCAAGGAGTCTCACTCCATGAGTCCCTGCT     636

    74 uSerProSerSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAspS     91
       ||||||||||||||||||||||||||||||||||||||||||||||||
   637 GAGTCCCAGCAGTGCTGCGAGGCCCTGAGCTGAATGGAAGAGACT          686

    91 erIleIleLysProValHisSerSerIleLeuGlyGlnGluPheCysPhe    107
       ||||||||||||||||||||||||||||||||||||||||||||||||
   687 CCATTATCAAGCCAGTACACAGCTCCATCCTGGGCCAGGAGTTCTGCTTT    736

   108 GluValThrThrSerSerGlyThrLysCysPheAlaCysArgSerAlaAl    124
       ||||||||||||||||||||||||||||||||||||||||||||||||
   737 GAGGTAACAACATCGTCTGGGACAAAAATGTTTGCCTGTCCGGTCTGCAGC    786

   124 aGluArgAspLysTrpIleGluAsnLeuGlnArgAlaValLysProAsnL    141
       ||||||||||||||||||||||||||||||||||||||||||||||||
```

us-09-294-298-6.rge

```
1687 AGAGCACCAGGCCAACCTGCGGATGTGCTGCGGAGTTGGCCCTGTCGCAAGG 1736

 441 alValAsnSerHisCysValPheProArgGluLeuLysGluValPheAla      457
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1737 TGGTCAACTCCCATTGCGTGTTCCCGAGGGAGCTGAAGGAGTGTTTGCA      1786

 458 SerTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAspAr     474
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1787 TCATGGCGGCTGCGCTGCGCAGAGCGGGGCCGGGAGGACATTGCTGACAG     1836

 474 gLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetS     491
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1837 GCTGATCAGCGCCTCGCTCTTCCTGCGCCGGCCATCATGT     1886

 491 erProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnThrSer     507
     ||||||||||||||||||||||||||||||||||||||||||||||||||
1887 CGCCCAGTCTGTTTGGACTGATGCAGGAGTACCCAGATGAGCAGACCTCA     1936

 508 ArgThrLeuThrLeuIleAlaLysValAlaIleGlnAsnLeuAlaAsnPheSe     524
     ||||||||||||||||||||||||||||||||||||||||||||||||||||
1937 CGAACCCTCACCCTCATCGCCAAGGTTATCCAGAACCTGGCCAACTTTTC     1986

 524 rLysPheThrSerLysGluAspSerPheLeuGlyPheMetAsnGluPheLeuG     541
     |||||||||||||||||||||||||||||||||||||||||||||||||||
1987 CAAGTTTACCTCAAAGGAGGACTTCCTGGGCTTCATGAACGAGTTTCTGG     2036

 541 luLeuGluTrpGlySerMetGlnGlnPheLeuGlnPheLeuGlnLeuSerAsnLeu     557
     |||||||||||||||||||||||||||||||||||||||||||||||||
2037 AGCTGGAGTGGGGTTCTATGCAGCAATTCTTGTATGAGATATCCAACCTG     2086

 558 AspThrLeuAsnSerSerPheGluGlyTyrIleAspLeuProGlyLeuGlyAr     574
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2087 GACACACTGAACAACAGCAGCAGTTTGAGGGCTACATAGACTTGGGCCG     2136

 574 gGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProGlnLeuS     591
     |||||||||||||||||||||||||||||||||||||||||||||||||
2137 CGAGCTCTCCACACTTCACGCCCTCTCTGGGAGGTGCTGCCCCAGCTCA     2186

 591 erLysGluAlaLeuLeuLysLeuGlyLeuProLeuProArgLeuLeuSerAsp     607
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2187 GCAAGGAGGCCCTCTGAAGCTGGGCCCGTCCCCGGCTCCTCAGCGAC     2236

 608 IleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnProSerArgGl     624
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2237 ATCAGCACAGCCCTGAGGAACCCTAACATTCAAAGGCAGCCGAGCCGCCA     2286

 624 nSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSerAlaG     641
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2287 GAGCGAGCGCGCTCGGTCTCAGCCCATGTGTCGCGGCGGCCGTCAGCCCG     2336

 641 luMetGlnGlyTyrMetMetLeuArgAspLeuAsnSerSerIleAspLeuGln     657
     |||||||||||||||||||||||||||||||||||||||||||||||||
2337 AGATGCAGGGCTACATGATGCGGGACCTCAACAGCTCCATCGACCTTCAG     2386

 658 SerPheMetAlaArgGlyLeuAsnSerSerMetAlaArgLeuPr     674
     |||||||||||||||||||||||||||||||||||||||||||
2387 TCCTTCATGGCTCGAGGCCTCAACAGCTCTATGGACAATGGCTCGCCTCCC     2436

 674 oSerProThrLysGluLeuProProProGlyGlyLeuGlyLysA     691
     |||||||||||||||||||||||||||||||||||||||||||
2437 CTCCCCAACCAAGGAGAAAACCCCCGGCCCCCTCCCCGTGGGGTAAAG     2486

 691 spLeuPheTyrValSerArgArgProLeuProLeuAlaArgSerProAlaTyr     707
     |||||||||||||||||||||||||||||||||||||||||||||||||
2487 ACCTGTTCTATGTGAGCCGGCCACCACTGCCCCCCGGCCCCCCCAGCATAC     2536

 708 CysThrSerSerAspIleThrGluProGluProGluGlnAsnLysMetLeuSerVa     724
     |||||||||||||||||||||||||||||||||||||||||||||
2537 TGCACGAGCAGCTCCGGACATCACAGAGCCGGAGCGAGATGCTGAGTGT     2586

 724 lAsnLysSerValSerMetLeuAspLeuGlnGlyLysGlyGlyLysGlyTyrA     741
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2587 CAACAAGAGTGTCTCCATGCTGGACCTGCAGGGCGACGGGCCTGGGGGCC     2636
```

```
 741 rgLeuAsnSerSerValSerAsnLeuAlaAlaValGlyAlaAspLeuLeu     757
     |||||||||||||||||||||||||||||||||||||||||||||||||
2637 GCCTTAACAGCAGTAGTGTTTCCAACCTGGCAGCTGTTGGGGACCTGTTG     2686

 758 HisSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaPr     774
     |||||||||||||||||||||||||||||||||||||||||||||||
2687 CACTCAAGCCAGGCTTCACTGACAGCAGCCTTGGGGTTGCGGCCTGCACC     2736

 774 oAlaGlyArgLeuSerGlnGlySerGlySerIleThrAlaAlaGlyM     791
     |||||||||||||||||||||||||||||||||||||||||||||||
2737 TGCCGGGCGCCCTCTCCCAAGGGAGTGGCTCTTCCATCACAGCAGCCGGCA     2786

 791 etArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnGln     807
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2787 TGCGCCCTTCCAGCATGGGTGTCACTACGGATGGTGTCCCCGGCCCAGCA     2836

 808 LeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaAs     824
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2837 CTCGCCATCCCTCTTCCTCCAGAACCCTCTCTTCCATATGGCCGCCAA     2886

 824 pGlyProGlyProAlaGlyHisGlyGlySerSerGlyHisGlyProP     841
     |||||||||||||||||||||||||||||||||||||||||||||||
2887 TGGACCAGGGCCCCCAGCAGGCCATGGAGGGAGCAGTGGCCATGGTCCAC     2936

 841 roSerSerHisHisHisHisHisHisHisHisArgGlyGlyGlu     857
     ||||||||||||||||||||||||||||||||||||||||||||
2937 CTTCCTCCCATCACCACCACCACCATCACCACCACCGGGGGGGAGAA     2986

 858 ProProGlyAspThrPheAlaProPheHisGlyTyrSerLysSerGluAs     874
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2987 CCCCCAGGGGACCACTTTTGCCCCGTTCCATGGCTATAGAAAGAGCGAGGA     3036

 874 pLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHisSerH     891
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3037 CCTCTCTACAGGGGTCCCTAAGCCCCCTGCGGCCTCCATCCTTCCACAGCC     3086

 891 iSerTyrSerAspGluPheGlyProSerGlyThrAspPheThrArgArg     907
     |||||||||||||||||||||||||||||||||||||||||||||||||
3087 ACAGCTACAGTGATGAGTTTGGACCCTCTGGTACTGATTTTACCCGTCGG     3136

 908 GlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSerProProGlnIl     924
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3137 CAGCTCTCACTTCAGGACAACCTACAGCACATGCTCTCCCCGCCCCAGAT     3186

 924 eThrIleGlyProGlnArgProAlaProSerGlyProGlyGlySerG     941
     |||||||||||||||||||||||||||||||||||||||||||||||
3187 CACCATCGGTCCCCAGAGGCCAGCTCCCTCAGGGCCAGGAGGGGCAGTG     3236

 941 lyGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProLeuGlnArg     957
     |||||||||||||||||||||||||||||||||||||||||||||||
3237 GTGGGGCAGTGTGTGGGGGCGTGGGGGCCAGCCCACCTCCCTTGCAGAGG     3286

 958 GlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysProArgProSe     974
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3287 GGCAAATCTCAGCAGTTGACAGTGTCTCCCCAGAAACCCCGGCCCGTC     3336

 974 rSerGlyAsnLeuLeuGlnSerProGluProSerTyrGlyProAlaArgP     991
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3337 CAGCCGGAACCTATTGCAGTCCCCGGAACCAAGTTATGGTCCTGCCCGTC     3386

 991 roArgGlnIleSerLeuSerLysGluGlySerIleGlyGlySerGlyGly     1007
     |||||||||||||||||||||||||||||||||||||||||||||||||
3387 CACGGCAACAGAGCCCTCAGCAAAGAGGGCAGCATTGGGGGCAGCGGGGGC     3436

1008 SerGlyGlyGlyGlyGlyGlyGlyGlySerIleThrLysGlnHi     1024
     |||||||||||||||||||||||||||||||||||||||||||||
3437 AGCGGTGGCGGAGGGGGTGGGGGGCCTCAAGCCCTCACCAAGCAGCA     3486

1024 sSerGlnThrProSerThrLeuAsnProThrMetProAlaSerGluArgT     1041
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3487 TTCCCAGACTCCATCACACGCTGAACCCCACCATGCCCGCCTCGGAGCGGGA     3536
```

```
1041  hrValAlaTrpValSerAsnMetProHisLeuSerAlaAspIleGluSer  1057
      |||||||||||||||||||||||||||||||||||||||||||||||||
3537  CTGTAGCCTGGGTGTCCAATATGCCTCACCTGTCCGCTGACATCGAGAGT  3586

1058  AlaHisIleGluArgGluGluTyrLysLeuLysGluTyrSerLysSerMe  1074
      |||||||||||||||||||||||||||||||||||||||||||||||||
3587  GCACCATTGAGCGGGAAGAGTACAAGCTGAAGGAGTACTCGAAGTCCAT   3636

1074  tAspGluSerArgLeuAspArgValLysGluLysGluLysTyrGluGluIleHiS  1091
      |||||||||||||||||||||||||||||||||||||||||||||||||
3637  GGACGAGAGCCGACTGGACAGGGTGAAGGAGTACGAGGAGGAGATCCACT  3686

1091  erLeuLysGluArgLeuHisMetSerAsnArgLysLeuGluGluTyrGlu  1107
      |||||||||||||||||||||||||||||||||||||||||||||||||
3687  CACTGAAGGAAAGGCTACACATGTCCAACCGGAAGCTGGAAGAGTACGAG  3736

1108  ArgArgLeuSerGlnGluGluGlnThrSerLysIleLeuMetGlnTy    1124
      |||||||||||||||||||||||||||||||||||||||||||||||||
3737  CGGAGGCTGCTGTCCCAGGAAGAGCAGACCAGCAAGATCCTGATGCAGTA  3786

1124  rGlnAlaArgLeuGlnSerGluLysSerArgLeuArgGlnGlnInValG  1141
      |||||||||||||||||||||||||||||||||||||||||||||||||
3787  CCAAGCCCGCTGAGCAGCGAGCAGAGCGCTTGAGCAGCAGCAGGTGG    3836

1141  luLysAspSerGlnIleLeuSerIleIleGlyArgLeuMetLeuValGlu  1157
      |||||||||||||||||||||||||||||||||||||||||||||||||
3837  AGAAGGACTCCCAGATCAAGAGCATCATTGGCAGGCTGATGCTGGTGGAG  3886

1158  GluGluLeuArgArgAspHisProAlaMetAlaGluProLeuProGluPr  1174
      |||||||||||||||||||||||||||||||||||||||||||||||||
3887  GAGGAGCTGCGCCGGGACCACCCCGCCATGGCTGAGCCGCTGCCTGAACC  3936

1174  oLysLysArgLeuAspAlaGlnArgGlySerPheProProTrpValG    1191
      |||||||||||||||||||||||||||||||||||||||||||||||||
3937  CAAGAAGAGGCTGCTCGACGCTCAGAGAGGCAGCTTCCCCCCTTGGTCC   3986

1191  lnGlnThrArgVal...ArgTrpProHisLeuGlyThrAlaTrpProProPro  1206
      |||||||||     ||||||||||||||||||||||||||||||||||
3987  AACAAACCCGGTGTGACGCTGGACGCTGCCCCACCTTGGAACGGCCCTGGCCCCCCC  4036

1207  GlnProHisProProHisProGlyCys  1215
      |||||||||||||||||||||||||||
4037  CAGCCCCACCCCCCCACCCCGGGTGC  4063
```

seq_name: gb_ro:AB016962

```
                     /organism="Rattus norvegicus"
                     /sub_species="Sprague Dawley"
                     /db_xref="taxon:10116"
     gene            1124..4624
                     /gene="synGAP-d"
     CDS             1124..4624
                     /gene="synGAP-d"
                     /note="synaptic ras GTPase activating protein"
                     /codon_start=1
                     /product="synGAP-d"
                     /protein_id="BAA74972.1"
                     /db_xref="GI:4417207"
                     /translation="MEYPPAAPFRPSQGFLSRRLKSSIKRTKSQPKLDRTSSFRQILP
                     RFRSADHDRARLMQSFKESHSHESLLSPSSAAEALELNLDEDSIIKPVHSSILGQEFC
                     FEVTTSSGTKCFACRSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELPPK
                     KRYCELCLDDMLYARTTSKPRSASGDfVFWGEHFEFNNLPAVRALRLHLYRDSDKKR
                     KKDKAGYVGLVTVPVATLAGRHFTEQWYPVTLPTSGSGGMGSGGGGSGGSGGGGKG
                     KGGCPAVRLKARYQTMSILPMELYKEFAEYVINHYRMLCAVLEPALNVKGKEEVASAL
                     VHIQSTGKAKDFLSDMAMSEVDRFMERRHLIFRENTLATKAIEEYMRLIGQKYLKDA
                     IGEFITRALYRSEENCEVDPFKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKE
                     VFASWRLRCAERGREDIADRLISASLRFLFCPAIMSPSLFGLMQEYPDEQTSRTLTL
                     IAKVIQNLANFSKFTSKEDFLGFMNEFLELEWGSMQQFLYEISNLDLTNSSSFEGYI
                     DLGRELSTLHALLWEVLPQLSKEALLKLGPLPRLLSDISTALRNPNIQRQPSRQSERA
                     RSQPMVLRGPSAEMQGYMRDLNSSIDLQSFMARGLNSSMDMARLSPTKEKPPPPPP
                     GGGKDLFYVSRPPLARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGDGPGGRLNS
                     SSVSNLAAVGDLLHSSQASLTAALGLRPAPAGRLSQGSGSSITAAGMRLSQMGVTTDG
                     VPAQQLRIPLSFQNPLFHMAADGPGPPAGHGGSSCHGPPSSHHHHHHHRGGEPPG
                     DTFAPFHGYSKSEDLSTGVPKPPAASILHSHSYDEFGPSGTDFTRRQLSLQDNLQHM
                     LSPPOITIGPQRPAPSGPGGGSGGGSGGGGGQPPLQRGRSQQLIVSAAQKPRPSSG
                     NLLQSPEPSYGPARPRQQSLSKEGSIGGSGSGSGGGGGLKPSITKQHSQTPSTLNPT
                     MPASERTVAWVSNMPHLSADIESAHIERREYKLKEYSSMDESRLDREYEEEIHSLKE
                     RLHMSNRKLEEYERRLLSQEEQTSKILMQYQARLEQSEKRLRQQQVEKDSQIKSIIGS
                     PSLQADAGGGGAAPGPPRHG"
```

```
BASE COUNT    1043 a   1477 c   1314 g    967 t
ORIGIN
```

```
   9  ProHisPheProPheAlaProGlnInGlyPheLeuSerArgArgLeuLy   25
      |||     ||||| |||::||||||||||||||||||||||||||||
1136  CCCGCTGCCCCTTCCGGCCCTCCAAGGGCTTCCTGAGCCGGAGGCTAAA  1185

  25  sSerSerIleLysArgThrLysSerGlnProLysLeuAspArgThrSers  42
      |||||||||||||||||||||||||||||||||||||||||||||||||
1186  AAGCTCCATCAAACGTACAAAGTCACAACCCAAACTTGACCGGACCAGCA  1235

  42  erPheArgGlnIleLeuProArgPheArgSerAlaAspHisAspArgAla  58
      |||||||||||||||||||||||||||||||||||||||||||||||||
1236  GCTTTCGACAGATCCTGCCCCGCTTCCGAAGTGCTGACCATGACCGGGCC  1285

  59  ArgLeuMetGlnSerPheLysGluSerHisSerHisGluSerLeuLeuSe  75
      |||||||||||||||||||||||||||||||||||||||||||||||||
1286  CGGCTGATGCAGAGCTTCAAGGAGTCACTCCATGAGTCCCTGCTGAG   1335

  75  rProSerSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAspSerI  92
      ||||||||||||||||||||||||||||||||||||||||||||||||
1336  TCCCAGCAGTGCTGCTGAGGCCCTGGAGCTCAACCTGGATGAAGACTCCA  1385

  92  leIleLysProValHisSerSerIleLeuGlyGlnGluPheCysPheGlu  108
      |||||||||||||||||||||||||||||||||||||||||||||||||
1386  TTATCAAGCCAGTACACAGCTCCATCCTGGGCCAGGAGTTCTGCTTTGAG  1435

 109  ValThrThrSerSerGlyThrLysCysPheAlaCysArgSerAlaAlaGl  125
      |||||||||||||||||||||||||||||||||||||||||||||||||
```

us-09-294-298-6.rge

```
  425 uHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysValV 442
      |||||||||||||||||||||||||||||||||||||||||||||||||||
 2386 GCACCAGGCCAACCTGCGGATGTGCTGTGAGTTGGCCCTGTGCAAGGTGG 2435

  442 aIAsnSerHisCysValPheProArgLeuLysGluValPheAlaSer 458
      ||||||||||||||||||||||||||||||||||||||||||||||
 2436 TCAACTCCCATTGCGTGTTCCCGAGGGAGCTGAAGGAGGTGTTTGCATCA 2485

  459 TrpArgLeuArgCysAlaGluArgGlyArgGluAlaAspIleAlaAspArgIe 475
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2486 TGGCGGCTGCGCTGTGCAGAGCGGGGCCGGGAGGCAATTCTGACAGGCT 2535

  475 uIleSerAlaSerLeuPheLeuCysProAlaIleMetSerP 492
      ||||||||||||||||||||||||||||||||||||||||||
 2536 GATCAGCGGCCGCCTTCCTCCGCGTTCCTCTCCGCCCATCATGTCGC 2585

  492 roSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnThrSerArg 508
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2586 CCAGTCTGTTTGGACTGATGCAGGAGTACCCAGATGAGCAGACCTCACGA 2635

  509 ThrThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerLy 525
      ||||||||||||||||||||||||||||||||||||||||||||||||
 2636 ACCCTCACCCTCATCGCCAAGGTTATCCAGAACCTGGCCAACTTTTCCAA 2685

  525 sPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGluL 542
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2686 GTTTACCTCAAAGGAGGACTTCCTGGGCTTCATGAACGAGTTTCTGGAGC 2735

  542 euGluTrpGlySerMetGlnGlnPheLeuTyrGluIleSerAsnLeuAsp 558
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2736 TGGAGTGGGGTTCTATGCAGCAATTCTTGTATGAGATATCCAACCTGGAC 2785

  559 ThrLeuThrAsnSerSerPheGluGlyTyrIleAspLeuGlyArgGl 575
      |||||||||||||||||||||||||||||||||||||||||||||||
 2786 ACACTGACCAACAGCAGCTTTGAGGGCTACATAGACTTGGGCCGCGGA 2835

  575 uLeuSerThrLeuHisAlaLeuTrpProGlnLeuSerL 592
      ||||||||||||||||||||||||||||||||||||||
 2836 GCTCTCCACACTTCACGCCTCTCTGGGAGTGCTGCCCCAGCTCAGCA 2885

  592 ysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSerAspIle 608
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2886 AGGAAGCCCTCCTGAAGCTGGGCCCGCTGCCCCGGCTCCTCAGCGACATC 2935

  609 SerThrAlaLeuArgAsnProAsnIleGlnArgGlnProSerArgGlnSe 625
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2936 AGCACAGCCCTGAGGAACCCTAACATCCAAAGGCAGCCGCCAGAG 2985

  625 rGluArgAlaArgSerGlnProMetValLeuArgGlyProSerAlaGluM 642
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2986 CGAGCGCGCTCCGGTCTCAGCCCATGGTGCTGCGCGGCCGTCAGCCGAGA 3035

  642 etGlnLeuTyrThrMetAlaArgLeuAsnSerSerIleAspLeuGlnSer 658
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 3036 TGCAGGGCTACATGATGCGGACCCTCAACAGCTCCATCGACCTTCAGTCC 3085

  659 PheMetAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSe 675
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 3086 TTCATGGCTCGAGGCCTCAACAGCTCTATGGACATGGCGCTCCCCTCCTC 3135

  675 rProThrLysGluLysProProProGlyGlyGlyGlyLysAspL 692
      ||||||||||||||||||||||||||||||||||||||||||||
 3136 CCCAACCAAGGAGAAACCCCCGCCCCCTCCGGTGGGGGGTAAAGACC 3185

  692 euPheTyrValSerArgProLeuAlaArgSerProAlaTyrCys 708
      |||||||||||||||||||||||||||||||||||||||||||||
 3186 TGTTCTATGTGAGCCGGCCACCACTGGCCCGCTCCCCAGCATACTGC 3235

  709 ThrSerSerSerAspIleThrGluProGluGlnLysMetLeuSerValAs 725
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 3236 ACGAGCAGCTCCGACATCACAGAGCCGGAGCAGAAGATGTCGAGTGTCAA 3285
```

```
 1436 GTAACAACATCGTCTGGGACAAAATGTTTGCCTGTCGGTCTGCAGCCGA 1485

  125 uArgAspLysTrpIleGluAsnLeuGlnArgAlaValLysProAsnLysA 142
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1486 AAGGGACAAATGGATTGAGAATCTACAGAGGGCTGTGAAACCCAACAAGG 1535

  142 sPAsnSerArgArgValAspAsnArgValLeuLysTrpIleIleGluAla 158
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1536 ACAACACCCGCGGGTAGATAACGTGCTGAAACTATGGATCATAGAAGCT 1585

  159 ArgGluLeuProProLysLysArgTyrTyrCysGluLeuCysLeuAspAs 175
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1586 CGAGGCTGCCCCCCAAGAAGCGATATTACTGCGAGTTATGCCTGGACGA 1635

  175 pMetLeuTyrAlaArgThrThrSerLysProArgSerAlaSerGlyAspT 192
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1636 CATGCTCTATGCACGGACCACTTCCAAGCCCCGCTCAGCCTCAGGAGACA 1685

  192 hrValPheTrpGlyGluHisPheGluAsnAsnLeuProAlaValArg 208
      ||||||||||||||||||||||||||||||||||||||||||||||
 1686 CTGTCTTTTGGGGAGCACTTCGAGTTTAACAACCTGCCTGCTGTCCGG 1735

  209 AlaLeuArgLeuHisLeuTyrArgAspSerArgLysLysLysAs 225
      |||||||||||||||||||||||||||||||||||||||||||||
 1736 GCGCTGCGGCTGCATCGTGACCGTGGACAAAAAGCGGAAGAAGGA 1785

  225 pLysAlaGlyTyrValGlyLeuValThrValProValAlaThrLeuAlaG 242
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1786 CAAGGCAGGCTACGTTGGCCTGGTGACTGTTCCAGTGGCCACCCTGGCTG 1835

  242 lyArgHisPheThrGluGlnTrpTyrProValThrLeuProThrGlySer 258
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1836 GGCGCCACTTCACAGAGCAGTGGTACCCCGTGACCCTGCCAACAGGAAGT 1885

  259 GlyGlySerGlyGlyMetGlySerGlyGlyGlyGlyGlyGlySerGlyGlyGl 275
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 1886 GGGGGCTCTGGGGGTATGGGCTCGGGGGGAGGGGGGGGTCAGGGGGCGG 1935

  275 yserGlyGlyGlyGlyGlyGlyCysProAlaValArgLeuLysAlaA 292
      |||||||||||||||||||||||||||||||||||||||||||||||
 1936 CTCAGGGGGCAAAGGGAAAGGAGGAGCGTGTCCTGGTGCGGCTGAAGGCC 1985

  292 rgTyrGlnThrMetSerIleLeuProMetGluLeuTyrTyrLysGluPheAla 308
      |||||||||||||||||||||||||||||||||||||||||||||||
 1986 GTTACCAGACAATGAGTAATCCTGCCCATGGAGCTATATAAGGAGTTTGCA 2035

  309 GluTyrValThrPheAsnHisTyrArgMetLeuCysAlaValLeuGluProAl 325
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2036 GAATATGTGACCAACCACTACCGCATGCTGTGTGCCGTGCTGGAGCCCGC 2085

  325 aLeuAsnValLysGlyLysGluGluGluValAlaSerAlaLeuValHisIleL 342
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2086 CCTCAATGTCAAGGGCAAGGGAGAGCCCAAGGAGCTCGCACTGGTTCACATCC 2135

  342 euGlnSerThrGlyLysLysAlaLysAspPheLeuSerAspAlaMetSer 358
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2136 TGCAAAGCACACAGAAGAAGGCCAAGGACTTCCTTTCAGACATGGCCATGTCA 2185

  359 GluValAspArgPheMetGluArgGluHisLeuIlePheArgGluAsnTh 375
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2186 GAGGTAGACCGGTTCATGGAGCGGGAACACCTCATATTCCGCGAGAACAC 2235

  375 rLeuAlaThrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysT 392
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2236 GCTCGCCACTAAAGCCATAGAAGAGTATATGAGGCTGATTGGCCAGAAAT 2285

  392 yrLeuLysAspAlaIleGlyLysGluPheIleArgAlaLeuTyrGluSerGlu 408
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2286 ACCTCAAGGATGCCATTGGGAAGGAGTTCATCCGGGCTCTGTATGAATCTGAG 2335

  409 GluAsnCysGluValAspProIleLeuCysThrAlaSerSerLeuAlaGl 425
      ||||||||||||||||||||||||||||||||||||||||||||||||||
 2336 GAGAACTGTGAAGTAGACCCCATCAAGTGCACAGCGTCCAGTCTGGCAGA 2385
```

```
 725 nLysSerValSerMetLeuAspLeuGlnGlyAspGlyProGlyGlyArgL  742
     |||||||||||||||||||||||||||||||||||||||||||||
3286 CAAGAGTGTCCATGACGCTGGACCTGCAGGGCGACGGGCCTGGGGGCCGCC 3335

 742 euAsnSerSerValSerValSerAsnLeuAlaAlaAlaValGlyAspLeuLeuHis  758
     ||||||||||||||||||||||||||||||||||||||
3336 TTAACAGCAGTAGTGTTCCAACCTGGCAGCTGTTGGGGACCTGTTGCAC 3385

 759 SerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaProAl  775
     ||||||||||||||||||||||||||||||||||||||
3386 TCAAGCCAGGCTTCACTGACAGCAGCCTTGGGGTTGCCGCCTGCCACCTGC 3435

 775 aGlyArgLeuSerGlnGlySerGlySerIleThrAlaAlaGlyMetA  792
     ||||||||||||||||||||||||||||||||||||||
3436 CGGGCGCCTCTCCCAAGGAGTGGCTCTTCCATCACAGCAGCCGGCCATGC 3485

 792 rgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnGlnLeu  808
     ||||||||||||||||||||||||||||||||||||||
3486 GCCTCAGCCAGATGGGTGTCACTACGGATGGTGTCCCCGCCCAGCAACTG 3535

 809 ArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaAspGl  825
     ||||||||||||||||||||||||||||||||||||||
3536 CGCATCCCCTTCCTTCCAGAACCCCTCTTCCATATGGCTGCCCGATGG 3585

 825 yProGlyProProAlaGlyHisHisGlyGlySerSerGlyHisGlyProPro  842
     ||||||||||||||||||||||||||||||||||||||
3586 ACCAGGGCCCCCGCAGGCCATGGGAGGCAGTGGCCATGGTCCACCTT 3635

 842 erSerHisHisHisHisHisHisHisHisHisHisArgHisArgProSer  858
     ||||||||||||||||||||||||||||||||||||||
3636 CCTCCCATCACCACCACCACCATCACCACCACGAGGGGGGGAGAACCC 3685

 859 ProGlyAspThrPheAlaProPheHisGlyTyrSerLysSerGluAspLe  875
     ||||||||||||||||||||||||||||||||||||||
3686 CCAGGGGACACTTTTGCCCCGTTCCATGGCTATAGCAAGAGCGAAGGACCT 3735

 875 uSerThrGlyValProLysProAlaAlaSerIleLeuHisSerSerHisS  892
     ||||||||||||||||||||||||||||||||||||||
3736 CTCTACAGGGGTCCCTAAGCCCCTGGGGCCCTCCATCTTCACAGCCACA 3785

 892 erTyrSerAspGluPheGlyProSerGlyThrAspPheThrArgArgGln  908
     ||||||||||||||||||||||||||||||||||||||
3786 GCTACAGTGATGAGTTTGGACCCCTCTGGTACTGATTTTACCCGTCGGCAG 3835

 909 LeuSerLeuGlnAspAsnLeuGlnHisMetLeuSerProProGlnIleLeuTh  925
     ||||||||||||||||||||||||||||||||||||||
3836 CTCTCACTTCAGGACAACAACCTACAGCACATGCTCTCCCCCCAGATCAC 3885

 925 rIleGlyProGlnArgProAlaProSerGlyProGlyGlyGlySerGlyG  942
     ||||||||||||||||||||||||||||||||||||||
3886 CATCGGTCCCCAGAGGCCCAGCTCCCTCCAGGGCCAGGAGGGGCAGTGGTG 3935

 942 lyGlySerGlyGlyGlyGlyGlyGlnProProLeuGlnAsnArgGly  958
     ||||||||||||||||||||||||||||||||||||||
3936 GGGGCAGTGGTGGGGGGCGGTGGGGGGGCCACCTCCCTTGCAGAGGGGGC 3985

 959 LysSerGlnGlnLeuThrValSerAlaAlaAlaGlnLysLysProSerSe  975
     ||||||||||||||||||||||||||||||||||||||
3986 AAATCTCAGCAGTTGACAGTGAGTGCTGCCCCAGAAACCCCGCCGTCCAG 4035

 975 rGlyAsnLeuLeuGlnSerProGluProSerTyrrGlyTyrProAlaArgProA  992
     ||||||||||||||||||||||||||||||||||||||
4036 CGGGAACCTATTGCAGTCCCCGGAACCAAGTTATGGYCCTGCCCGTCCAC 4085

 992 rgGlnGlnSerLeuSerLysGluGlyLysSerIleIleGlyGlySerGlyGlySer 1008
     ||||||||||||||||||||||||||||||||||||||
4086 GGCAACAGCAGCCCTCAGCAAGAGGGCAGCATTGGGGCAGCAGGGGGCAGC 4135

1009 GlyGlyGlyGlyGlyGlyLeuGlyGlyLeuLysLysProSerIleThrHisLysGlnHisSe 1025
     ||||||||||||||||||||||||||||||||||||||
4136 GGTGGCGGAGGGGGTGGGGGGGCTCAAGCCCTCCATCACCAAGCAGCATTC 4185

1025 ↑GlnThrProSerThrLeuAsnProThrMetProAlaSerGluArgThrV 1042
```

```
     |||||||||||||||||||||||||||||||||||||||||||||||||
4186 CCAGACTCCATCCACGCTGAACCCCGATGCCGGCCTCGGAGCGGACTG 4235

1042 alAlaTrpValSerAsnMetProHisLeuSerAlaIleGluSerArgIle 1058
     |||||||||||||||||||||||||||||||||||||||||||||
4236 TAGCCTGGGTGTCCAATATGCCTCACCTGTCCGCTGAACYCGAGAGTGCA 4285

1059 HisIleGluArgGluGluTyrLysLeuLysGluLysTyrSerLysSerMetAs 1075
     |||||||||||||||||||||||||||||||||||||||||||||
4286 CACATTGAGCGGGAAGAGTACAAGCTGAAGGAGTACTCCAAGTCCATGGA 4335

1075 pGluSerArgLeuAspArgValLysGluTyrGluGluIleHisSerL 1092
     |||||||||||||||         ||||||||||||||||||||||
4336 CGAGAGCCGACTGGACAGG......GAGTACGAGGAGGAGATCCACTCAC 4379

1092 euLysGluArgLeuHisMetSerArgAsnArgLysLeuGluGluTyrGluArg 1108
     |||||||||||||||||||||||||||||||||||||||||||||
4380 TGAAGGAAAGGCTACACATGTCCAACCGGAAGCTGGAAGAGTACGAGCGG 4429

1109 ArgLeuLeuSerGlnGluGlnGlnThrSerLysIleLeuMetGlnTyrGl 1125
     |||||||||||||||||||||||||||||||||||||||||||||
4430 AGGCTGCTGTCCCAGGAAGAGCAGCAGACCAGCAAGATCCTGATGCAGTACCA 4479

1125 nAlaArgLeuGlnLnSerGluLuLysArgLeuArgGlnGlnValGluL 1142
     |||||||||||||||||||||||||||||||||||||||||||||
4480 AGCCCGCCTGGAGCAGGAGCAGGAAGAGCGTTGAGGCAGCAGCAGTGGAGA 4529

1142 ysAspSerGlnIleLeuLysSerIleIleIleGly..........Arg.LeuMe 1154
     |||||||||||||||||||||||||||||           ::  ||||
4530 AGGACTCCCAGATCAAGAGCATCATTGGCAGCCCGTCCCTTCAGGCTGAT 4579

1154 tLeuValGluGluGluLeuArgArgAspHisProAlaMetAlaGluProL 1171
     |||||||||||||||||||||||||||||||||||||||||||||
4580 GCTGGTGGAGGAGGAGCTGCGCCGGGACCACCCCGCCATGGCTGAGCCGC 4629

1171 euProGluProLysLysArgArgLeuLeuAspAlaGln.ArgGlySerPhePr 1187
     |||||||||||||||||||||||||||||| ||||||||||||||
4630 TGCCTGAACCCAAGAAGAGGAGGCTGCTCGACGCTCAGGAGAGGCAGCTTCCC 4679

1187 oProTrpValGlnGlnThrArgVal...ArgTrpProHisLeuGlyThrA 1203
     ||||||||||||||||||||   ||||||||||||||||||||||
4680 CCCTTGGGTCCACAAACCCGGTGTGACGCTGGCCCACCTTGGAACGG 4729

1203 laTrpProProGlnProHisProProHisProGlyCysArgSerGlnArg 1219
     |||||||||||||||||||||||||||||||||||||||||||||
4730 CCTGGCCCCCCCAGCCCCACCCCCACCCCCGGCTGCAGATCACAGAGA 4779

1220 ThrAlaSerSerGlyThrPro 1226
     |||||||||||||||||||||
4780 ACGGCGAGTTCCGGAACACCG 4800
```

seq_name: gb_htg20:AL161903

```
seq_documentation_block:
LOCUS        AL161903    150956 bp    DNA              HTG       29-SEP-2000
DEFINITION   Homo sapiens chromosome 6 clone RP11-175A4, *** SEQUENCING IN
             PROGRESS ***, 3 unordered pieces.
ACCESSION    AL161903
VERSION      AL161903.12    GI:10039702
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 150956)
AUTHORS      Tracey,A.
TITLE        Direct Submission
JOURNAL      Submitted (28-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
             CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
             requests: clonerequests@sanger.ac.uk
COMMENT      On Sep 9, 2000 this sequence version replaced gi:9796876.
             --------    Genome Center
             Center: Sanger Centre
```

```
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-------- Project Information
Center project name: bA175A4
-------- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; L08752; 99% of reads
Chemistry: Dye-primer-amersham; 0% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 145680 bases at least Q40
Consensus quality: 147560 bases at least Q30
Consensus quality: 148600 bases at least Q20
Insert size: 150756; sum-of-contigs
Insert size: 148326; 10.8% error; agarose-fp
Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality
coverage: 6.54x in Q20 bases; agarose-fp
--------
*  NOTE: This is a 'working draft' sequence. It currently
*  consists of 3 contigs. The true order of the pieces
*  is not known and their order in this sequence record is
*  arbitrary. Gaps between the contigs are represented as
*  runs of N, but the exact sizes of the gaps are unknown.
*  This record will be updated with the finished sequence
*  as soon as it is available and the accession number will
*  be preserved.
*       1    54611: contig of 54611 bp in length
*   54612 54711: gap of     100 bp in length
*   54712   70293: contig of 15582 bp in length
*   70294 70393: gap of     100 bp in length
*   70394  150956: contig of 80563 bp in length.
FEATURES        Location/Qualifiers
source          1..150956
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="6"
                /clone="RP11-175A4"
                /clone_lib="RPCI-11.1"
misc_feature    1..54611
                /note="assembly_fragment:01454
                clone_end:T7
                vector_side:left"
                54712..70293
                /note="assembly_fragment:02542
                fragment_chain:1"
misc_feature    70394..150956
                /note="assembly_fragment:02844
                fragment_chain:1
                clone_end:SP6
                vector_side:right"
BASE COUNT  39575 a  37381 c  36411 g  37389 t     200 others
ORIGIN

alignment_scores:
     Quality: 4060.50          Length:      3203
       Ratio: 3.882              Gaps:        19
Percent Similarity: 32.657      Percent Identity: 31.876

alignment_block:
US-09-294-298-6 x AL161903/rev  .

Align seg 1/1 to reverse of: AL161903  from: 1   to: 150956

    16 GlnGlyPheLeuSerArgArgLeuLysSerSerIleLysSerArgThrLy 32
       ||| ::       :::|||:::   |||:::::|   |||||  ||
107723 CAGAGCAAGACTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAGAG 107674

    32 sSerGlnProLysLeuAspSerAlaThrSerSerPheArgGlnIleLeuProA 49
       ::::||||||         ::: :::|:::||||:|:|||||
107673 TGGGAAGCCC........TGATCCCTTCCTCCTGGAACCTGCCTGCCTG 107633
```

```
    49 rgPheArgSerAlaAspHisAspArgAlaArgLeuMetGlnSerPheLys 65
       |||||||||||||||||||||||||||||||
                    AGGGCCGGCTGATGCAAAGCTTTAAG 107604
    66 GluSerHisSerGluSerLeuLeuSerProSerSerAlaAlaGluAl 82
       |||||||||||||||||||||||||||||||||||||||
107632 CC............GAGTCACACTCTCATGAGTCCTTGCTGAGTCCAGTGCAGCTGAGGC 107554
    82 aLeuGluLeuAsnLeuAspGluAspSerIleIleLysProValHisSerS 99
       |||||||||||||||||||||||||||||||||||||||||||||||
107603 GAGTCACACTCTCATGAGTCCTTGCTGAGTCCAGTGCAGCTGAGGC 107504
    99 erIleLeuGlyGlnGluPheCysPheGlu.Val.......... 109
       |||||||||||||||||||||  ::
107553 ATTGGAGCTGAATGGAGATTCCATTATCAAGCCAGTGCACAGCT 107454
   109 ......................................... 109
107503 CCATCCTGGGCCAGGAGTTCTGTTTTGAGGTACTGGGTCTGTGGGCTGG 107404
   109 GGAGGGCCAAAGGACAGGGGTGATGGAAGGTGGGGGGCCAGGAGGTCTAG 107354
       .................................................
107453 AGAAAGTGGCACAGTGGGACATCAGGAAACAGAAACTTCTGGGACTG 107304
   109 ......................................... 109
107403 GAGGAAAGGTAGGCCAGGGAGGAAGAGAAGGTAGCAGAGTCTCCTCCCC 107254
   110 .........................ThrThrSer 112
                               |||:|||||
107353 TCTGTAGCCCTTTCCTCAACTCCACACTCCTTTCTAGGTAACAACTTCA 107204
   113 SerGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAspLysTr 129
       |||||||||||||||||||||||||||||||||||||||||||
107303 TCAGGAACAAAATGCTTTGCCTGTCGGTCTGCGGCGAAAGAGACAAATG 107154
   129 pIleGluAsnLeuGlnArgAlaValLysProAsnLys. 141
       ||||||||||||||||||||||||||
107253 GATTGAGAATCTGCAGCGGGCAGTAAAGCCCAACAAGGTATTGGGAATA 107104
   141 .............................................. 141
107203 AAGGGGACACAACCTGTGTAGGCAAAGGTTGCACCAACACGGGAAGCTG 107054
   141 .............................................. 141
107153 GGGGGCCTAGGGAGGAAAGTGAGTTAAAGGAGGAGGCTTGGGGAAGGAG 107004
   141 .............................................. 141
107103 AGGATTGGAGGTACAGTGTATCTGGACAAGCAGGGGGAGACCCCCATTATT 106954
   141 .............................................. 141
107053 CTGAGTCCCCCATTCTTTTCGCTTTCTGTACTGCTACCCTGCCTTACGA 106904
   141 .............................................. 141
107003 TCTCTTCCCTGCCATAGAGTCATAGAGTCATAGAGTTACAGAGTTGATGGGCTCT 106854
   141 .............................................. 141
106953 GGAAATTCTTTAGTCTAGCTTCCCTGCAGCAGGAATGCTTCACTAATAC 106804
   141 .............................................. 141
106903 CCCGCAGGACACATCAAATACACTTAGCCAAGTTTCTGTACCTTGGTTTC 106754
   141 .............................................. 141
106853 CTCTCTAAGACAGAGGGAACTGGTTGGCAATGGGTTGATTAAATGATCTC 106804
   141 .............................................. 141
106803 ........................................ 141
```

```
106753  TAAGGTCCCTTGGCACATAAATTCTATGAGTCTGTCTTCCCAAAACTT  106704
141     ...............................................
105853  CAGGGAATCCTCCTTGCCTCCTTCCTAGCTTCCGGAGGTTGCCAGCAGTC  105804
141     ...............................................

106703  CATGCTTCCAGTTCGTCAGCTATTCTCTGTATCCCAGTTTCTAGCCAAGC  106654
141     ...............................................
105803  CTTGGCCATTCCTGGGCTTATAACTGCATCCGTCTAATTCTGCCTCCATC  105754
141     ...............................................

106653  TTACAGTCCTAGTCACACTCCTCTGGGGGAACTCTTGTGTCTTGATGTC  106604
141     ...............................................
105753  TTCATGTAGCTGGCTTCCTTCTCTGTGTGTCTCTGTATCCTGTATCCTGTG  105704
141     ...............................................

106603  CTTAAAGAACTCAACCCAAGATGCTGACATGATCTGACCTATGCAGAGTA  106554
141     ...............................................
105703  TCTCCAAATCTCCCTCTCCATATAAAGACACCAGTTTTATAAGGTGGGT  105654
141     ...............................................

106553  CAAACACCATGTTCCCTTCAACACACTGCAATATCACCTGAGCCTTACC  106504
141     ...............................................
105653  TAAGGGTCCACTCTAATTCAGTATGGCCTCTTCTTTTTTGAGATGGAATAT  105604
141     ...............................................

106503  ATTAGCTTGTGTCTCAATTGTATCCCCCAGGCCCTTACTGATTCTTTTT  106454
141     ...............................................
105603  CGCTCTTGTTCCCCAGGCTTGGAGTGCAATGGCATGAGCTCAGCTCACTG  105554
141     ...............................................

106453  TTTTTTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAG  106404
141     ...............................................
105553  CAACCTCCGCCCTTCCAGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAG  105504
141     ...............................................

106403  TGGCGTGATCTTGGCTCACTGCAACCTCCGCCTCCTGGGTTTAAGCAATT  106354
141     ...............................................
105503  TAGCTGGATTACAGGCACGTGCCACTATGCCCAGCTAATTTGTTTTTGT  105454
141     ...............................................

106353  CCTGTGCCTCAGCCTCCCCAGTAGCTGGGATTACAGGCCATGCACCACCAC  106304
141     ...............................................
105453  ATTTTTATTAGAGACGGGGTTTCACTATGTTGGCCAGGCTGGCTCGAACT  105404
141     ...............................................

106303  ACCCGACTAATTTTGTATTTTTAGTAGAGACCGAGGTTTCACCATGTTGG  106254
141     ...............................................
105403  CCTGACCTCAGGCAATCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTA  105354
141     ...............................................

106253  CCAGGCTAGTCTCAAACTCCTGACCTCAGGTGTCTGCCTGCCTCAGCCT  106204
141     ...............................................
105353  CAGGAATGAGATACCATGCTGGCCTCCTTCTTCTTAGTTGATTACATTGCA  105304
141     ...............................................

106203  CCCAAAGTGCTAGGATTACAGGCATGAGCCACCACATCTGGCCAAGCTTT  106154
141     ...............................................
105303  GGGATCCTGTTTCCAAATAAGATCACATTCACAGGTTCAGGGTAGACATG  105254
141     ...............................................

106153  ACCCATTCTATATGCAATTCTTTTTCACAATTTCTGATAGTCTCTGCAG  106104
141     ...............................................
105253  AGTTTTGGTGGGATACTATTCAACCCAGTACACCGTCTCACCTTATGATT  105204
141     ...............................................

106103  GACTTTCCAGTTCCCTTCAAATTCTACATTAATATTTTGGCTTGTTATT  106054
141     ...............................................
105203  GCTGTCTGTATCCCACTTTCTTACTGTCTCCTCTTTTGGTGTCTCT  105154
142     .........................................AspAsnSe  144
                                                 |||||||
105153  GTCCTCTCCCCTTAACATCATGCCACCCACCATGCCAGGACAACAG  105104

106053  CCAGCTTTGAAATCTTTCCAATACTGTTGTTGTTGTCTAGTGTATGTCT  106004
141     ...............................................
144     rArgArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGluL  161
        |||||||||||||||||||||||||||||||||||||||||||||||||
105103  CCGCCGGGTAGACAATGTGCTAAAGCTGTGGATCATAGAAGCCCCGGAGC  105054

106003  ACGCTTCTACAGTAGCTTCCTGGGGCTGCTATAACAAACTGCTACAAACT  105954
141     ...............................................
161     euProProLysArgTyrTyrCysGluLeuCysLysAspMetLeu  177
        ||||||||||||||||||||||||||||||||||||||||||||
105053  TGCCCCCAAGAAGCCGTACTACTGTGAGCTCTGCCTGATGACATGCTG  105004

105953  GGGTCACGTAAAACAACAGAAATTTATTCTCTTCACAGTTCAGGAGGCTGG  105904
141     ...............................................
178     TyrAlaArgThrThrSerLysProArgSerAlaSerGlyAspThrValPh  194
        |||||||||||||||||||||||||||||||||||||||||||||||
105003  TATGCACGCACCACCTCCAAGCCCGCTCCGCCTCTGGGGACACCGTCTT  104954

105903  AAGTCCAAAAGCAAGGTATCAGCAGGGCCACCGCTCTCTTTGATGCTGTAG  105854
141     ...............................................
194     eTrpGlyGluHisPheGluAsnAsnLeuProAlaValArgAlaLeuA  211
        |||||||||||||||||||||||||||||||||||||||||||||||
104953  CTGGGGCGAGCACTTCGAGTTTAACAACCTGCCCGCGTCCCGTGCCCCTGC  104904
```

us-09-294-298-6.rge

Fri Jan 19 06:20:05 2001

```
 398  yGluPheIleArgAlaLeuTyrGluSerGluGluAsnCysGluValAspP  415
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104003 AGAATTCATCCGTCCTGCTTCGTATGAATCTGGAGAAAACTGCGAGGTAGACC 103954

 415  rolleLysCysThrAlaSerSerLeuAlaThrLeuAlaGlyArgHi  431
      |||||||||||||||||||||||||||||||||||||||||||||||
103953 CTATCAAGTGCACAGCATCCAGTTTGGCAGAGCACCAGGCCAACCTGCGA 103904

 432  MetCysCysGluLeuAlaLeuCysLysLysValValAsn..........  443
      |||||||||||||||||||||||||||||
103903 ATGTGCTGTGAGTTGGCCCTGTGCAAGGTGGTCAACTCCCACTGGTGAGA 103854

 443  ........................................  443
103853 CTTGGAACGCTGGGCTGGGGGGCCAGGGTCGGGGGGAATTATGTGTTCATC 103804

 443  ........................................  443
103803 TGTTCATCTATCTGTCCATCCTCAAAGAGGACTGAGCACCATTTATGGGC 103754

 443  ........................................  443
103753 AAAGCATTGTTCTAGGCGCTATAGAGCAAACAGGTGAAAGAGGCCTGGTC 103704

 443  ........................................  443
103703 CCTGCCCTCAGAGGGCCTCCACCAGAGGCCATAATGGGACAAATTAGAAGAAAAAA 103654

 443  ........................................  443
103653 AAAAAAAGCCACAGAGCCATAATGTGTGTAAGGTGCTGAGTAAGGGTCC 103604

 443  ........................................  443
103603 CCCAACCTCTGTGTGACATAAGGTCAGAGAGAGGCAGAGCTTTGAGAT 103554

 443  ........................................  443
103553 AAGTGGGAAGAGGTGCCCCCTTGGGTAGGCTTTGAAGACTGGTTTAGGT 103504

 443  ........................................  443
103503 TCTGATATATGGACATAGTTGGCAAGAAAGACATTTCAGAGAAGAGGCTGT 103454

 443  ........................................  443
103453 GAGAAAGGCACATGTGTGATCGTGAAAAGGCCCAGGAGTTTTCAGGGGAC 103404

 443  ........................................  443
103403 ATTAAAGTAGGTTAGTAGCAATTACATCAGGTTTAGTGGAGCATGTGCCT 103354

 443  ........................................  443
103353 CATAATGGGAGGTGGCGGGAGAGATGTCTGGGCAGCAGGAAGAATTAGCTTTAG 103304

 443  ........................................  443
103303 AAACTGAAGGCCTCAAGGAGTCTGAGGTCATTGTGTAGGCCTTGGGATCC 103254

 443  ........................................  443
103253 CATTAAAGGTGTCAGAAATATTGTGATTTAGAAGATTAATCCTAGGCTG 103204

 444  .......SerHis.Cys..............  446
                ||||| |||
103203 GGCACGCGGTGGCTCACACCTGTAATCCAGCACTTTGGGAGGCTGAGGCGG 103154

 446  ........................................  446
103153 GCAGACCACCTGAGGTCAGGAGTTTGAGACCAGGCTGACCAACATGGAGA 103104

 446  ........................................  446
103053                                              103104
```

```
 211  rgLeuHisLeuTyrArgAspSerAspLysLysArgLysAspLysAla  227
      |||||||||||||||||||||||||||||||||||||||||||||||
104903 GGCTGCATCTGTACCGTGACTCAGACAAAAAGCGCAAGAAGGACAAGGCA 104854

 228  GlyTyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgHi  244
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104853 GGCTATGTCGGCCTGGTGACTGTGCCAGTGGCCACCCTGGCTGGGCGCCA 104804

 244  sPheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGlyS  261
      |||||||||||||||||||||||||||||||||||||||||||||||||
104803 CTTCACAGAGCAGTGGTACCCTGTAACCCTGCCAACAGGCAGTGGGGGAT 104754

 261  erGlyGlyMetGlySerGlyGlyGlySerGlyGlyGlySerGlySerGly  277
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104753 CTGGGGGCATGGGTGGGAGGGGGGAGGGGGTGGCTTCAGGG 104704

 278  GlyLysGlyLysGlyGlyCysProAlaValArgLeuLysAlaArgTyrGl  294
      |||||||||||||||||||||||||||||||||||||||||||||||||
104703 GGCAAGGGCAAAGGAGGTTGCCCCGGCTGTGCGGCTGAAAGCACGTTACCA 104654

 294  nThrMetSerIleLeuProMetGluLeuTyrLysGluPheGluAlaGluTyrV  311
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104653 GACAATGAGCACTTTGCCCATGGAGCTCTATATAAAGAGTTTGCAGAGTATG 104603

 311  alThrAsnHisTyrArgMetLeuCysAlaValLeuGluHisIleLeuGlnAsn  327
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104603 TCACCAACCATTATCGGATGCTGTGTGCAGTCTTGGAGCCCGCCCTGAAT 104554

 328  ValLysGlyLysGluGluValAlaAlaSerAlaLeuHisIleLeuGlnSe  344
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104553 GTCAAAGGCAAGGAGGAGGTTGCCAGTGCACTAGTTCACATCCTGCAGAG 104504

 344  rThrGlyLysAla.Lys..............................  349
      |||||||||| :::
104503 TACAGGCAAGGCCAAGGTGAGTGTTGTGCCCTCAGGAAAGTGACTGG 104454

 349  ........................................  349
104453 GAATGGGCACTTGCTTGGGGGTTAGTGAGGACAGGGCAAATTCACGAGAT 104404

 350  ..................................AspPheL  352
                                          ||||||
104403 TGGGTGTGCAGAGCTGACACTTGGATTTTCCTGGGCCTCAGGAGACTTCC 104354

 352  euSerAspMetAlaMetSerGluValAspArgPheMetGluArgGluHis  368
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104353 TTTCAGACAATGGCCATGTCCGAGGTCAGACCGGTTCATGGAACGGGAGCAC 104304

 369  LeuIlePheArgGluAsnThrLeuAlaThrLysAlaIleGluGluTyrMe  385
      ||||||||||||||||||||||||||||||||||||||||||||||||||
104303 CTCATATTCCGCGAGAACACGCTTGCCACTAAAGCCATAGAAGAGTATAT 104254

 385  tArgLeuIleGlyGlnLysTyrLeuLysLysAspAla.Ile........  397
      ||||||||||||||||||||||||||||| :::
104253 GAGACTGATTGGTCAGAAATACCCCAAGGATGCCATTGGTATGGCCCACA 104204

 397  ........................................  397
104203 CTTCAGGCGCCCTCTTCTTCCCAAACCTGCCCAGAGTGTCACCCAGACCCCAA 104154

 397  ........................................  397
104153 GTCCACCCCTTCCACAGCTTGATACTTCCTAACCCAGTCCTAGGACTCC 104104

 397  ........................................  397
104103 AGCCTCCAACACCTGATTCTGAAATTCCCCAACCCTGCCACCCCCCTTC 104054

 398  ...............................Gl  398
                                     ||
104053 CCTGCCCTTGGAAAGTGTGACCACCACACCCCTCTTGTGCCCCCACCCCCAGG 104004
```

us-09-294-298-6.rge

```
103103 AACCCTGCTCTCTACTAAAAATACAAAATTAGCCAGATGTGGTGGCACATG 103054      102203 CCCTCCTTCCCTGTCTCTTCACCCCTGTTTCACACCCTCACCTCCTAC 102154
446 .................................................... 446           446 ..................................................... 446
103053 CCTGTAATCCCAGCTATTCGGGAGGCAGGAGGAATCACTTGAATC 103004             102153 CACCCCCTCAGCCATGTTCCCTGGAAGCTGAGGGTCTCGGGGCTCAGTC 102104
446 .................................................... 446           446 ..................................................... 446
103003 TGGGAGGTGGAGTTGTAGTGAGCCGAGATCACACCATTGCCACTCTAGCC 102954      102103 CCGGTCTCTCTTCTCTCTCTCGTCTCCCCACCCTTCCCC 102054
446 .................................................... 446           447 ....ValPheProArgGluLeuLysGluValPheAlaSerTrpArgLeu 461
102953 TGGGCAACAAGAGTGAAACTCCGTCTCAAAAAAAAAAATTAATCCATAA 102904        102053 CCAGCGGTGTTCCCGAGGCAGCTGAGGGAGGTGTTTGCTTCATGCGGCTG 102004
446 .................................................... 446           462 ArgCysAlaGluArgGlyArgGluAspIleAlaSerArgLeuIleSerAl 478
102903 GTAGAGTGTATGTAGTAGAGTAGAGGGATGTATGTTGGGGATGACTGGA 102854     102003 CGCTGCCAGAGCGAGCCCGGAGGACATCGCAGACAGGCTTATCAGCGCC 101954
446 .................................................... 446           478 aSerLeuPheLeuArgPheLeuCysProAlaIleMetSerProSerLeuP 495
102853 ATAGAGCCAGCTGGGGCTCGAGGGCCATGAAAGTCAGATCCTGAATCAGA 102804    101953 CTCACTCTTCCTGGCTTCCTCCTCGCCCAGCGATTATGTCGCCCAGTCTCT 101904
446 .................................................... 446           495 heGlyLeuMetGlnGluTyrProAspGluGlnThrSerArgThrLeuThr 511
102803 ACAGTGACAGAAGTTAGGAAGGAGCTACTGGAAGGACCCTCGGGAAAG 102754       101903 TTGGGCTTATGCAGGAGTACCCAGATGAGCAGACCCTCACGAACCCTCACC 101854
446 .................................................... 446           512 LeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerLys.......... 525
102753 GATCAGTAGGAATTGTCAGCTTATTGGAAAGGAGGAGCATGTCTGTGGG 102704       101853 CTCATTGCCAAGTCATCCAGAACCTGGCCAACTTTTCCAAGTGAGGGAA 101804
446 .................................................... 446           525 .................................................... 525
102703 GAGTCACGGATGACTGACTGATCTCAAGAGCCATGAGGCTGGTGGTTGGGAGACCGGT 102654  101803 GCTTCAGGAGTGGGCAGGGCAGGGAGTGGCAGGGAGTGGCAGGGC 101754
446 .................................................... 446           525 .................................................... 525
102653 GGGCTCATTGACAACCAGGAAAGTCAGGAGGAGGAGCCAGTGCCAGTGT 102604      101753 TGGGGTCGGCAAGAAAGGGTCTCCCAGTCCCCAGAGATCCTGAGATGGG 101704
446 .................................................... 446           525 .................................................... 525
102603 GGGGGCCGATGGTGAGCTCTGCTTTACACCAGCCGAGTTTAAGGTGTCAGT 102554   101703 GAGGCTATGATACCTTGTCGTGTGTGTATGTGTGTATGTGTGTGTGT 101654
446 .................................................... 446           526 .................................................... P 526
102553 GGGACATTGAAGTGGAACTGGGAGGGTGCGGGGAGTGAGCTCTGCAGCCATT 102504  101653 GTGTGTGTGTGTGTGTATGTGACCTTTATCTCTCCATTCTTGGCTAGCGT 101604
446 .................................................... 446           526 heThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGluLeu 542
102503 CCAGGGACTGGGGATCATGCCTGGGGCACCTCCATCCCCATTCCCCTGGA 102454    101603 TTACCTCAAAGGAGCTTTCTGGGCTTCCATGAATGAGTTTCTGGAGCTG 101554
446 .................................................... 446           543 GluTrpGlySerMetGlnGlnPheLeuTyrGluIleSerAsnLeuAspTh 559
102453 ATCCAGAGAGAGTTGGGGGTCCGAGCTCCTGTGTACCTCAAGTGACCCTCC 102404   101553 GAATGGGGTTCCATGCAGCAGTTTTTGTATGAGATCTCCAATCTGGACAC 101504
446 .................................................... 446           559 rLeuThrAsnSerSerPheGluGluTyrIleAspLeuGlyTyrIleArgGluL 576
102403 ATCTCTCTCCCATCTCTCTCTCCCGTGTGTCTGTTTTTCTTCCTCC 102354         101503 GCTAACCAACAGCAGTAGCTTTGAGGGTTACATCGACTTGGGCCGAGAGC 101454
446 .................................................... 446           576 euSerThrLeuHisAlaLeuLeuTrpGluValLeuProGlnLeuSerLys 592
102353 TCTCCTTGTCTCTCTCCCACACCCCTTCTTCTCCCCACGTGTCTCTC 102304       101453 TCTCCACACTGCCATGCCCTACTCTGGAGGTGCTGCCCCAGCTCAGCAAG 101404
446 .................................................... 446           593 ..Glu............................................... 593
102303 CCCTTCACCTTCTCTCCCCCATTCTCTCTCCCTAATCTGTCTGTTCC 102254        101403 GTCAGCAGATCCCCTCTTTGCCCTATCCCCAGATGGCTCCAGAGGTTCCT 101354
446 .................................................... 446           593 .................................................... 593
102253 CTCTGCCCATGGCCCCCTTCTTCAAGCAGCCTCCCCATCTTGCTCCTGCGGT 102204   101353 GGAGCCTGAGAAACTACCCTTGAAGATTTTTTTTTCTCCCCTTGTTTCTC 101304
446 ..;                                                             593 .................................................... 593
                                                                      101303 GAGGTGTCACCACTACTATCCCAACTCAGCCCCTCCACCTGCACCCTC 101254
```

us-09-294-298-6.rge

```
593   .........................................................                        593
                                                              SerSerIleAspLeuGlnSerPhe  659
                                                              ||||||||||||||||||||||||
101253 AGAGGCCCTCTTAGAGCTGGGCACTGAGCCCCCAGGTAACAGCCTCACCC  101204  GCTCGCCCCTCTTCCATCTCTCAGCTCCATCGACCTTCAGTCCTTC  100304

594   ........AlaLeuLeuLysLeuProLeuProArgLeuLeuSerAs         607   MetAlaArgGlyLeuAsn........................  665
              ||||||||||||||||||||||||||||||||:::||                 ||||||||||||||
101203 TTCCAGGAAGCCCCTCGAAGCTGGGTCCACTGCCCCGCTCCTCAACGA  101154  ATGGCTCGAGGCCTCAACAGTGAGGGCTCTCCCCTCCCCGCGCCTCCT  100254

607   pIleSerThrAlaLeuArgAsnProAsnIleArgGlnArgGlnSerArgG  624   .........................................  665
       |||||||||||||||||||||||||||||||||||||||
101153 CATCAGCACAGCTCTGAGGACCCCAACATCCAAAGGCAGCCAAGCCGCC  101104  CTCCTCCTCCTGTATGTTCCCTCCCACTCCACTGGCCTTCGCCCTACTC  100204

624   lnSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSerAla  640   .........................................  665
       |||||||||||    ||| |||:::::||||||||||||||
101103 AGAGTGAGCGGCCGCCCAGCCTGTGTACTGCGGGGGCCATCGGCT  101054  CTCTCCTCTCCTCCTCCATGGACCTCATCTCTCCATATGCCCAGCCC  100154

641   GluMetGlnGlyTyrMetMetArgAspLeuAsn....................... 651   .........................................  665
       ||||||||||||||||||||||||||||||||||||
101053 GAGATGCAGGGCTACATGATGCGGGACCTCAACAGGTGAGCACCCTGGGA  101004  TGCCCCATCCCTTCTCTTGCTGCCCCATCTCCCCTCCCTCTAGGCCTCA  100104

651   ....................................................  651   .........................................  665
101003 CAGCCAAGGCCTGTGCCCTAGGAGCCCTTCCTCTATTCTAGATACTCCTCA  100954  CCCCCTTCCGGAGGGGCCCTGTCCTTTCCCTTTACTCACCTGTCCCCTC  100054

651   ....................................................  651   .........................................  665
100953 CTGGGGCCCCACATGGCATCTCTCTAGGGCTTGAAAGAAGGGAGGAAAAAGC  100904  CCATCCTCCCTGCCTCCCTCTTCAGGGCTGCCACCCGCTAGCTCTCAGCC  100004

651   ....................................................  651   .........................................  665
100903 ACCAAGTTCTCAGGGGAGACGATAAGGAGACAGGTACAGTCAGTGGTAGG  100854  CTTCCCTCTGGGTCCCACTTTTCACCCCAAGGCCTGTGCCAGACCACAGC  99954

651   ....................................................  651   .........................................  665
100853 CTGAGAGCCCCTTTACAGCCCTGAGGGAGTGAGAGATGTGAGCTCTAGGAA  100804  AAGGTTCAATTGCTAGGAGCCCTGACCTTACCTTCTGCTTGTGTGCCCCC  99904

651   .........................................  666   Ser.SerMetAspMetAlaArgLeuProSerProThrLy  678
                                                       :::::||||||||||||||||||||||||||
100803 TAGGGCTGAGGCTCCACCAACTCACGGCTTAGTTGTAAGCCTAGAGCATC  100754  TTCCCTTCTGACAGCTCTATGGACATGGCTCGCCTCCCTCCCCAACCAA  99854

651   ....................................................  678   sGluLysProProProProGlyGlyLysAspLeuPheTyrV  695
       |||||||||||||||||||||||||||||||||||||||||
100753 CCTGCTGCAAGCTCTGATTTGCTGCGTCCCTCTGCCTGCCCATGCTAGTCCC  100704  GGAAAAGCCACCCCACCACCGCCCGTGGTGGTAAAGACCTGTTCTATG  99804

651   ....................................................  695   alSerArgProLeuAlaArgSerProAlaTyrCysThrSerSer  711
       |||||||||||||||||||||||||||||||||||||||||||
100703 CAGGCTGAGGTTCAGCCAGCATGTCATGTCAGCCATGTGTCAAAATGTTC  100654  TAAGCCGTCCACCCTGCCCGTCCTCACCAGCCATACTGCCACGAGCAGC  99754

651   ....................................................  712   SerAspIleThrGluProGluGlnLysMetLeuSerValAsnLysSerVa  728
       |||||||||||||||||||||||||||||||||||||||||||||||||
100653 AAACATCTCAGTAATAGCTAGTGAATAAGCACTTCCCCAGCCCCCGACC  100604  TCGACAATCACAGCACCCAGAGAAGATGCTGAGTGTCAACAAGAGTGT  99704

651   ....................................................  728   lSerMetLeuAspLeuGlnLysGlnValAspGlyProGlyGlyArgLeuAsnSerS  745
       ||||||||||||||||||||||||||||||||||||||||||||||||
100603 ACAACCCCACAGACCTCCCCATGATCCAGCCACTTAGAGCAGTGACAGCAG  100554  GTCCATGCTGGACTTACAGGGTGATGGGCCCGGCCCCCTCAACAGCA  99654

651   ....................................................  745   erSerValSerAsnLeuAlaAlaValGlyAspLeuHisSerSerGln  761
       |||||||||||||||||||||||||||||||||||||||||||||
100553 AAGCCTAGCAGGGCCTGCAGCCTGCTCCAGAGTCCCAGCCTCCATTCTGA  100504  GCAGTGTTTGGAACCTGGCGCGCCTAGGGGACCTGCACTCAAGCCAG  99604

651   ....................................................  762   AlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaProAlaGlyArgLe  778
       |||||||||||||||||||||||||||||||||||||||||||||
100503 TAGGTGGTGCCCGTGCCCTGCTTCATTTGCTGCTCATTATTTGATGATGGGCCCT  100454  GCCTCGCTGACAGCAGCCTTGGGGCTACAGGCCTGCGCCGACGCGCCT  99554

651   ....................................................  778   uSerGlnGlySerGlySerSerIleThrAlaAlaGlyMetArgLeuSerG  795
       |||||||||||||||||||||||||||||||||||||||||||||
100453 GCTTTTTTCCTCTGCCCGTGTCTTCCTCCATGACACCATCCCACC  100404  CTCCCAGGGCAGTGGCTCATCCATCACGGCCGCCATGCGCCTCAGCC  99504

651   ....................................................  795   lnMetGlyValThrThrAspGlyValProGlnGlnLeuArgIlePro  811
       |||||||||||||||||||||||||||||||||||||||||||||
100403 ATTCCCGCCTCTCCTTTCATTTGTCCACATCTCTCTTCTCTGTCTGT  100354  AGATGGGTGTCACCACAGACGGTGTCCCTGCCAGCAACTGCGAATCCCC  99454

                                                             812   LeuSerPheGlnAsnProLeuPheHisMetAlaAlaAspGlyProGlyPr  828
```

Fri Jan 19 06:20:05 2001

```
99453 |||||||||||||||||||||||||||||||||||||||||||||||||
      CTCTCCTTCCAGAACCCTCTTCTCCACATGGCTGCTGATGGCCAGGTCC 99404
 828  oProAlaGlyHisGlySerGlyHisGlyProProSerSerHisH 845
      ||||:::||||||||||::::|||||||||||||||||||||
99403 CCCAGGCGGCCATGGAGGGGCCGGCCACCTTCCTCCCATC 99354
 845  isHisHisHisHisHisHisArgGlyGlyGlyHisGlyGluProProGlyAsp 861
      ||||||||||||||||||||||||||||||||||||||||||||||||
99353 ACCACCACCACCATCACCACCACCACCGAGGTGGAGCCCCCTGGGGAC 99304
 862  ThrPheAlaProPheHisGlyTyrSerLysSerGluAspLeuSerThrGl 878
      ||||||||||||||||||||||||::|||||||||||||::::|
99303 ACCTTTGCCCCATTCCATGGCTATAGCAAGAGTGAGGACCTTCTTCCGG 99254
 878  yValProLysProProAlaAlaSerIleLeuHisSerHisSerTyrSerA 895
      |||||||||||||||||||||||||||||||||||||||||||||||
99253 GGTCCCCAAGCCCCCTGCCTCCATCCTTCATAGCCACAGCTACAGTG 99204
 895  spGluPheGlyTyrProSerGlyThrAspPheThrArgArgGlnLeuSerLeu 911
      |||||||||||||||||||||||||||||||||||||||||||||||||
99203 ATGAGTTTGGACCCTCTGGCACTGACTTCACCCGTCGGCAGCTTTCACTC 99154
 912  GlnAspAsnLeuGlnHisMetLeuSerProProGlnIleThrIleGlyPr 928
      ||||||||||||||||||||||||||||||||||||::::||
99153 CAGGACAACCTGCAGCACATGCTGTCCCCTCCCCAGATCACCATTGGTCC 99104
 928  oGlnArgProAlaProSerGlyProGlyGlySerGlyGlyGlySerGlySerG 945
      |||||||||||||||||||||||||||||||||||||||||||||||
99103 CCAGAGGCCAGCCCCCTCCAGGGCCTGGAGGGTGGGAGCGGTGGGGGCAGCCG 99054
 945  lyGlyGlyGlyGlyGlnProProLeuGlnArgGlyGlyLysSerGlyGln 961
      ||||||||||||||||||||||||||||||||||||||||||||||
99053 GTGGGGGGTGGCGGGGGGCCAGCCCCTTCCATTCGAGAGGGGGCAAGTCTCAG 99004
 962  GlnLeuThrValSerAlaAlaGlnLysProArgProSerSerGlyAsnLe 978
      |||||||||||||||||||||||||||||||||||||||||||||
99003 CAGTTGACAGTCAGCGCAGCCCAGAAAACCCCGGCCATCCAGCCGGGAATCT 98954
 978  uLeuGlnSerProGluProSerTyrGlyProAlaArgProArgGlnGlnS 995
      |||||||||||||||||||||||||||||||||||||||||||||
98953 ATTGCAGTCCCCAGAGCCAAGTTATGGCCCCCGCCCCGTCCCACGGCAACAGA 98904
 995  erLeuSerLysGluGlyLysSererIleGlyGlyLySerGlyGlySerGlyGly 1011
      ||||||||||||||||||||||||||||||||||||||||||||
98903 GCCTCAGCAGAGGAGGCAGCCATTGGGGGCAGCGGGGGCAGCGGTGGCGGA 98854
1012  GlyGlyGlyGlyLeuLysLysProSerIleThrLys............... 1022
      |||||||||||||||||||||||
98853 GGGGGCGTGGGGGGCTGAAGCCCCTGCAACCACCACCAGCAGTAGGTGAAGGCAG 98804
1022  ................................................... 1022
98803 GAGGAAGGCGGCGGCTGGGGTCACACAGGAGGGGAAGAGGAGATGGGGGGT 98754
1022  ................................................... 1022
98753 GGGGTTGAAACAGAGTCTGTGGCCTGAAGTTACAAATCTTCTTGCCTCCTT 98704
1022  ................................................... 1022
98703 TTGGCCATTAAACAAATGTGTATAGAGGGGCTGCTATGTGCCATGTGCTA 98654
1022  ................................................... 1022
98653 TGCCAGGCACTAAGGATACGGCACTGAACCCTCTTAGCACTCTCATTCCA 98604
1022  ................................................... 1022
98603 GAGGGGGATTAATCCATAAGTAGAAATGGGGGTGACTGGAATAGGGCCAGCT 98554
1022  .,................................................. 1022
```

```
98553 GGGGGCTTGGAGGCATCAAGAGATCAGATCCTCGAATCAGAACAACAGGAGAA 98504
1022  ....................................................... 1022
98503 ATTCTCTGGCATTCAACTCACATCTCTGGCATTCAGAGGTGATAAAGGCT 98454
1022  ................................................... 1022
98453 ACAGCAGGCAATAGGACTGGGACTGTCGTCCTTTGGCTGTGCTGTGTTG 98404
1022  ................................................... 1022
98403 CCCTCTAAATGTACCCTGCTTTTCCCCACCTTTCTTTCTTCTGTTCGCCC 98354
1023  ................GlnHisSerGlnThrProSerThrLeuAsnProT 1034
      ||||||||||||||||||||||::||:|
98353 TCACTGTGCCTTGTCCCAGCATTCTCAGACACCATCCACATTGAACCCCA 98304
1034  hrMetProAlaSerGluArgThrValAlaTrpValAlaSerAsnMetProHis 1050
      |||||||||||||||||||||||||||||||||||||||||||||||||||
98303 CAATGCCAGCCTCTGAGCGGACAGTGGCCTGGGTCTCCAACATGCCTCAC 98254
1051  LeuSerAlaAspIleGluSerAlaHisIleGluArgGluGluTyrLysLeu 1067
      ||||||||||||||||||||||||||||||||||||||||||||||||||
98253 CTGTCGGCTGACATCGAGAGTGCCCACATCGAGCGGGAAGAGTACAAGCT 98204
1067  uLysGluTyrSerLysSerMetAspGluSerArgLeuAspAspArgValLys 1083
      ||||||||||||||||||||||||||||||||||||||||||::::
98203 CAAGGAGTACTCAAAATCGATGGAATGGAGAGCGCCGGCTGGATAGGGTACGA 98155
```

seq_name: gb_pr7:HSDJ570F3

```
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP4-570F3 is
            from the library RPCI-4 constructed at the Roswell Park Cancer
            Institute by the group of Pieter de Jong. For further details see
            http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            IMPORTANT: This sequence is not the entire insert of clone
            RP4-570F3 It may be shorter because we only sequence overlapping
            sections once, or longer because we arrange for a small overlap
            between neighbouring submissions.
            The true left end of clone RP4-570F3 is at 1 in this sequence. The
            true left end of clone ICRF6c-CB2046 is at 94671 in this sequence.
            The true left end of clone ICRF6c-CK0721Q is at 20826 in this
            sequence. The true right end of clone ICRF6c-CK0721Q is at 61562 in
            this sequence.
FEATURES          Location/Qualifiers
source            1..94770
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="6"
                  /clone_lib="RPCI-4"
                  /clone="RP4-570F3"
mRNA              complement(join(1012..1185,1669..2740,3135..3176,
                  3870..4048,4255..4456,4664..4900,6710..6854,7066..7210,
                  7338..7961,10016..10114,10324..10477,12824..12945,
                  13378..13469,19727..19832,22032..22153,25299..25366))
                  /gene="dJ570F3.1"
                  /note="match: cDNAs: Em:AF048976 Em:AF058789 Em:AF058790
                  Em:AB016962 Em:AF050183 Em:AF047711 Em:AF053938"
                  /evidence=not_experimental
                  /product="dJ570F3.1 (homolog of the rat synaptic ras
                  GTPase-activating protein p135 SynGAP)"
gene              complement(1012..25366)
                  /gene="dJ570F3.1"
CDS               complement(join(<1012..1185,1669..2740,3135..3176,
                  3870..4048,4255..4456,4664..4900,6710..6854,7066..7210,
                  7338..7961,10016..10114,10324..10477,12824..12945,
                  13378..13469,19727..19832,22032..22153,25299..25320))
                  /gene="dJ570F3.1"
                  /note="match: proteins: Tr:O95174 Tr:O44242 Wp:CE23473
                  Wp:CE23471 Wp:CE23475 Wp:CE23472 Wp:CE23470 Wp:CE23474
                  Wp:CE23476"
                  /codon_start=1
                  /evidence=not_experimental
                  /product="dJ570F3.1 (homolog of the rat synaptic ras
                  GTPase-activating protein p135 SynGAP)"
                  /protein_id="CAB63778.1"
                  /db_xref="GI:6624587"
                  /translation="MSTAPFRDVRGPSMHRTQYVHSPYDRPGWNPRFCIISGNQLLML
                  DEDEIHPLLIRDRRSESSRNKLLRRTVSVPVEGRPHGEHEYHLGRSRRKSVPGGKQYS
                  MEGAPAAPFRPSQGFLSRRLKSSIKRTKSQPKLDRTSSFPRQILPRFRSADHDRARLMQ
                  SFKFSHSHESLLSPSSAAEALELNLDEDSIIKPVHSSILGQEFCFEVTTSSGTKCFAC
                  RSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELPKKRYCELCLDDMLY
                  ARTTSKPRSASGDTVFWGEHFEFNNLPAVRALRLHLYRDSDKRKRKDKAGYVGLTVVP
                  VATLAGRHFTEQWYPVTLPTCSGSGGMGSGGGGSGGSGGKGKGGCPAVRLKARYQ
                  TMSLLPMELYKEFAEYVTNHYRMLCAVLEPALNVKGKEEVASALVHILQSTGKAKDFL
                  SDMAMSEVDRFMEREHLIFRENTLATKAIEYMRLIGQKYLKDAIGEFIRALYESEEN
                  CEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKEVFASWRLRCAERGR
                  EDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTLTLIAKVIQNLANFSKF
                  TSKEDFLGFMNEFLELELMGSWMQQFLYEISNLDTLTNSSSPEGYIDLGRELSTLHALLW
                  EVLPQLSKEALLKLGPLPRLLNDISTALRNPNIORQPSRQSERRPQPVVLRGPSAEM
                  QGYMMRDLNSSIDLQSFMARGLNSSMDMARLPSPKPKPPPPGGGDLYVYSRPPL
                  ARSSPAYCTSSSDITEPEQKMLSVNKVSMLDLQGDPGGRLNSSVSNLAAVGDLLH
                  SSQASLTAALGLRPAPAGRLSQGSGSSITAAGMRLSQMGVTTDGVPAQQLRIPLSFQN
                  PLFHMAADGPGPGHGGGGHGPPGSHHHHHHHHRGGEPPGDTFAPFHGYKSKSED
                  LSSGVPKPPAASILHSHSYSDEFGPSGTDFTRRQLSLQDNLQHMLSPPQITIGPQRPA
                  PSGPGGSGGGSGGGGGGQPPLQRGKSQQLTVSAAQKPRPSSGNLLQSPEPSYGPAR
                  PRQQSLKEGSIGGSGGGGGLKPSITKQHSQTPSTLNPTMPASERTVAWVSNM
                  PHLSADIESAHIEREEYKLKEYSKSMDESRLDR"
repeat_region     2178..2210
                  /note="11 copies 3 mer ggt 90 conserved"
repeat_region     2945..3122
                  /note="89 copies 2 mer gg 56 conserved"
misc_feature      complement(4211..4599)
                  /gene="dJ570F3.1"
                  /note="match: GSS: Em:B68206"
repeat_region     4484..4539
                  /note="28 copies 2 mer ca 94 conserved"
repeat_region     4605..4649
                  /note="3 copies 15 mer 91 conserved"
repeat_region     4903..4970
                  /note="34 copies 2 mer ga 70 conserved"
repeat_region     4996..5190
                  /note="13 copies 15 mer 59 conserved"
misc_feature      complement(6393..6846)
                  /gene="dJ570F3.1"
                  /note="match: GSS: Em:AQ475907"
repeat_region     8077..8178
                  /note="MLT1C repeat: matches 6..109 of consensus"
repeat_region     8476..8841
                  /note="MLT1C repeat: matches 109..461 of consensus"
repeat_region     10989..11119
                  /note="FLAM_C repeat: matches 3..131 of consensus"
repeat_region     13130..13161
                  /note="16 copies 2 mer ca 93 conserved"
repeat_region     13256..13303
                  /note="24 copies 2 mer ct 75 conserved"
misc_feature      17111..17815
                  /note="CpG island"
                  /evidence=not_experimental
repeat_region     17369..17458
                  /note="6 copies 15 mer 70 conserved"
repeat_region     17370..17513
                  /note="72 copies 2 mer cc 63 conserved"
repeat_region     17515..17658
                  /note="72 copies 2 mer gg 59 conserved"
repeat_region     17631..17726
                  /note="32 copies 3 mer gcc 63 conserved"
repeat_region     19005..19078
                  /note="37 copies 2 mer ga 67 conserved"
repeat_region     19612..19691
                  /note="40 copies 2 mer cc 63 conserved"
repeat_region     19858..19911
                  /note="27 copies 2 mer cc 70 conserved"
repeat_region     24653..24808
                  /note="78 copies 2 mer cc 57 conserved"
repeat_region     24706..24831
                  /note="14 copies 9 mer ctcctccct 61 conserved"
repeat_region     25104..25289
                  /note="93 copies 2 mer cc 59 conserved"
repeat_region     25187..25294
                  /note="12 copies 9 mer ctcccccc 64 conserved"
repeat_region     25367..25465
                  /note="11 copies 9 mer ggggaggag 69 conserved"
repeat_region     25373..25522
                  /note="10 copies 15 mer 64 conserved"
repeat_region     25411..25469
                  /note="19 copies 3 mer agg 86 conserved"
misc_feature      26820..27191
                  /note="match: GSS: Em:AQ605113"
misc_feature      27145..27728
                  /note="CpG island"
                  /evidence=not_experimental
misc_feature      complement(27412..27910)
                  /note="match: GSS: Em:AQ369330"
mRNA              join(27713..28149,28320..28379,28488..28533,28661..28710,
                  28854..29088)
                  /gene="cICK0721Q.5"
                  /note="match: cDNAs: Em:AF106943 Em:E06811
                  match: ESTs: Em:W75983 Em:AA187766 Em:AI004315 Em:R77008
                  Em:N94792 Em:AA843316 Em:AA384803 Em:AI245618 Em:AI921804
                  Em:AI792323 Em:AI802586 Em:AI524460 Em:AA843230
                  Em:AA887246 Em:AI096824 Em:AI929640 Em:N32014 Em:AA725127
                  Em:AA357305 Em:AA218432 Em:AI332801 Em:AA090206
```

Fri Jan 19 06:20:05 2001

us-09-294-298-6.rge

```
          Em:AI050991  Em:AF034177  Em:AI168128  Em:AA622278  Em:AA069440
          Em:AI087362  Em:AI168128  Em:AI340323  Em:AI028045
          Em:AA934024  Em:AI073391  Em:AA331045  Em:AA708233
          Em:AI342398  Em:AA056366  Em:AI582796  Em:AI762297
          Em:AI340198  Em:AI541387  Em:N34313   Em:AI567726  Em:AI364919
          Em:AI310946  Em:AA181976  Em:AI042256  Em:R81592   Em:T49588
          Em:AI929571  Em:D52353   Em:AA035663  Em:N48534   Em:AI373814
          Em:AA971499  Em:H25294   Em:AA703121  Em:AA44626   Em:AA379516
          Em:AI313425  Em:W70287   Em:AA937820  Em:H21228   Em:AA471384
          Em:N71594   Em:AA095718  Em:AA377858  Em:H25235   Em:H52756
          Em:D52210   Em:AA329903"
          /evidence=not_experimental
          /product="dJ570F3.2 (POLYPEPTIDE FROM PATENTED CDNA
          EMBL:E06811)"

alignment_scores:
        Quality:  4059.50        Length:    2324
        Ratio:    4.314          Gaps:        12
Percent Similarity:  40.491    Percent Identity:  39.974

alignment_block:
US-09-294-298-6 x HSDJ570F3/rev   .

Align seg 1/1  to reverse of: HSDJ570F3   from:  1  to: 94770


 139 ProAsnLysAspAsnSerArgArgValAspAsnValLeuLysLeuTrpPiI  155
     ||| :::||||||||||||||||||||||||||||||||| |||||||||
7970 CCATGCCAGGACAACAGCCCGGCCGGGTAGACAATTGCTAAAGCTGTGGAT  7921

 155 eIleGluAlaArgGluLeuProProLysLysArgTyrTyrCysGluLeuC  172
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7920 CATAGAGGCCCGGGAGCTGCCCCCAGAAGCGGTACTACTGTGAGCTCT  7871

 172 ysLeuAspAspMetLeuTyrAlaArgThrThrSerLysProArgSerAla  188
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7870 GCCTGGATGACAATGTGTATGCAGCCACCACCCCAAGCCCCGCTCTGCC  7821

 189 SerGlyAspThrValPheThrpGlyLeuHisPheGluPheAsnAsnLeuPr  205
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7820 TCTGGGGACACGGTCTTCTTGGGGCGGAGCACTTCGAGTTAACAACCTGCC  7771

 205 oAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysA  222
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7770 GGCTGTCCGTGCCCTGCGGCTGCATCTGTACCGTGACTCAGACAAAAAGC  7721

 222 rgLysLysAspLysAlaGlyTyrValGlyLeuValThrValProValAla  238
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7720 GCAAGAAGACAAGGCAGGCAGGCTATGTCGGCCTGGTCACTGTGCCAGTTGGCC  7671

 239 ThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeuPr  255
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7670 ACCCTGGCTGGCCGCCACTTCACAGAGCAGTGGTACCCTGTAACCCTGCC  7621

 255 oThrGlySerGlyLySerGlyGlyMetGlySerGlyGlyLySGlyGlyLys  272
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7620 AACAGGCAGTGGGGGATCTGGGGGCATGGGGTTCGGGGAGGGGGGAGGGGGCT  7571

 272 erGlyGlySerGlyLysGlyLyGlyGlyGlyLysGlyGlyCysProAlaValArg  288
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7570 CGGGGGGGTGGCTCAGGGGCAAGGGCAAAGGAGGTTGCCCGGCTGTGCGG  7521

 289 LeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLy  305
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7520 CTGAAAGCACACGTTACCAGACAATGAGCATCTTGCCCATGGAGCTATATAA  7471

 305 sGluPheAlaGluThrValThrhrAsnHisTyrArgMetLeuCysAlaValL  322
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7470 AGAGTTTGCAGAGTATGTCACCAACCATTATCGGATGCTGTGTGCAGTCT  7421

 322 euGluProAlaLeuAsnValLysGlyLysGluGluValAlaAlaSerAlaLeu  338
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7420 TGGAGCCCGCCCTGAATGTCAAAGGAGTTGCCCAGTGCACTA  7371
```

```
 339 ValHisIleLeuGlnSerThrGlyLysAla Lys...............  349
     ||||||||||||||||||||||||  :::
7370 GTTCACATCCTGCAGAGTACAGGCAAGGCCAAGGTGAGTGTGTGCCCTC  7321

 349 ...............................................  349
7320 AGGGAAAGGTGACTTGGGAATGGCACTTGCTTGCGGGGTTAGTGAGGACA  7271

 349 ...............................................  349
7270 GGGCAAATTCACGAGATTGGGTTGTGCAGAGGCTGACACTGGATTTCC  7221

 350 ............AspPheLeuSerAspMetAlaMetSerGluValAspArgP  363
                 |||||||||||||||||||||||||||||||||||
7220 TGGGCCTCAGGACTTCCTTCAGACATGGCCATGTCTGAGGTAGACCGGT  7171

 363 heMetGluArgGluHisLeuPheArgGluAsnThrLeuAlaThrLys  379
     ||||||||||||||||||||||||||||||||||||||||||||||
7170 TCATGGAACGGGAGCACCTCATATTCCGGGAACACGCTTGCCACTAAA  7121

 380 AlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLysAspAl  396
     ||||||||||||||||||||||||||||||||||||||||||||||||||
7120 GCCATAGAGAGTATATGAGACTGATTGGTCAGAAATACCTCAAGGATGC  7071

 396 a.Ile...............................................  397
     |.:::
7070 CATTGGTATGGCCACACTCAGGCCCTCTTCTTCCCAAACCTGCCAGATG  7021

 397 ...............................................  397
7020 TCCACCCCAGACACCCCAAGTCCACCCTTCCACAGCTGATACTTCCTAACC  6971

 397 ...............................................  397
6970 CAGAGTCCTAGGACTCCAGCCTCCAACACCTGATTCTGAAATTCCCCAA  6921

 397 ...............................................  397
6920 CCCTGGCCACCCCTTCCCTGCCCTTGGAAAGTGCACCACCCTCTTG  6871

 398 ...........GlyGluPheIleArgAlaLeuTyrGluSerGluGl  409
                ||||||||||||||||||||||||||||||||||||
6870 TGCCCCCACCCCCAGGAGAATTCATCCGTGCTCGTGTATGAATCTGAGGA  6821

 409 uAsnCysGluValAspProIleLysCysThrAlaSerLeuAlaGluH  426
     ||||||||||||||||||||||||||||||||||||||||||||||||||
6820 AAACTGCGAGGTAGACCCTATCAAGTGCACAGCATCAGTTGGCAGAGC  6771

 426 isGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysValVal  442
     ||||||||||||||||||||||||||||||||||||||||||||||||||
6770 ACCAGGCCAACCTGCGAATGTGCTGTGAGTTGGCCCTGCCAAGGTGGTC  6721

 443 AsnSerHis..............................................  445
     ||||||||||
6720 AACTCCCACTGGTGAGACTGGGAACGCTGGGCTGGGGGGCCAGGGTCGGG  6671

 445 ...............................................  445
6670 GGAATTATGTGTTCATCTGTTCATCTATCTGTCCATCCTCAAAGAGGACT  6621

 445 ...............................................  445
6620 GAGCCACCATTTATGGGCAAAGCATTGTTCTAGGCGCTATAGAGCAAACAG  6571

 445 ...............................................  445
6570 GTGAAAGAGGGCCTGGTCTCCCTGCCCTCAGAGGGGGCCTTCCACCAGAATGGGGA  6521

 445 ...............................................  445
6520 CAAATTAGAAGAAGAAAAAAAAAAAGCCACAGAGCCCATAATGGTCGTGTAAG  6471

 445  445
```

Fri Jan 19 06:20:05 2001

```
6470 TGCTGAGTAAGGGTCCCCCACCTCTGTGTGACATAAGGTCAGAGAGAA 6421
 445 ................................................  445
6420 GGCAGCAGCTTTGAGATAAGTGGGGAAGAGGTGCCCCCTTGGGTAGGCTTT 6371
 445 ................................................  445
6370 GAAGACTGGTTTAGGTTCTGATATATGGACATAGTTGGCAAGAAAGACAT 6321
 445 ................................................  445
6320 TTCAGAAGAAGGCTGTGAGAAAGGCACATGTGTGATGGTGAAAAGGCCCA 6271
 445 ................................................  445
6270 GGAGTTTTCAGGGGACATTAAAGTAGGTTAGTAGCAATTACATCAGGTTT 6221
 445 ................................................  445
6220 AGTGGAGCATGTGCCTCATAATGGGAGTGGCGGGAGAGATGTCTGGGCA 6171
 445 ................................................  445
6170 GGAAGATTAGCTTTAGAAACTGGAAGGCCTCAAGGAGTCTGAGGTCATTG 6121
 445 ................................................  445
6120 GTAGGCCTTGGGATGCCATTAAAGGTGTCAGAAAATATTGTGATTTAGAAG 6071
 445 ................................................  445
6070 ATTAATCCATAGGCTGGGCACGGTGGCTCACACCTGTAATACCAGCACTT 6021
 445 ................................................  445
6020 TGGGAGGCTGAGGCGGGCGGAGACCACCTGAGGTTCAGGAGTTTGAGACCAGG 5971
 445 ................................................  445
5970 CTGACCAACATGGGAGAAACCCTGTCTCTCTACTAAAAATACAAAAATTAGCCA 5921
 445 ................................................  445
5920 GATGTGGTGGCGCACATGCCTGTAATCCCAGCTATTCGGGAGGCTAAGGCAG 5871
 445 ................................................  445
5870 GAGAATCACACTTGAATCTGGGAGGTGGGAGGTTGTAGTGAGCCGAGATCACA 5821
 445 ................................................  445
5820 CGATTGCCACTCCTAGCCTGGGCAACAAGAGTGAAACTCCATCTCAAAAAAA 5771
 445 ................................................  445
5770 AAAATTAATCCATAAGTAGAGTGTATGTGTAGTAGAGGGATGTATG 5721
 445 ................................................  445
5720 TTGGGGATGACTGGAAATAGAGCCAGCTGGGGGCTCGGAGGCATGAAAGT 5671
 445 ................................................  445
5670 CAGATCCTGAATCAGAACAGTGACACAGAGTTAGGAAGGAGCTACTGGAAG 5621
 445 ................................................  445
5620 GACCCCTCTCGGGAAAGGATCAGTAGGAGAATTGTCAGCTTATTGGAAGGAG 5571
 445 ................................................  445
```

```
5570 GGAGCATGTCTGTGGGAGTCAACGTCCATGAGCTGG 5521
 445 ...............................  445
5520 TGGTTGGGAGACCGGTGCCTCATTGACAACCAGGAAAGTCAGGAGGAGG 5471
 445 ...............................  445
5470 AGCCAGTTGGAGGTTGGGGGCGATGTGAGCTCTTTACCACCAGCCG 5421
 445 ...............................  445
5420 AGTTTAAGGTGTCAGTGGGACATTGAAGTGGAACTGGGAGGTGCGGGGAG 5371
 445 ...............................  445
5370 TGAGCTCTGAGCCATTCCAGGGACTGGGGATCATGCCTGGGCACCTCCA 5321
 445 ...............................  445
5320 TCCCCATTTCCCTGGAATCCAGAAGAGTTGGGGGTCCGAGCTCCCTGTA 5271
 445 ...............................  445
5270 CCTCAAGTGACCCTCCATCTCTCCCCATCTCTGTCTCTCCCTGGTGTCT 5221
 445 ...............................  445
5220 GTTTTTCTTCTCTCTCTCTTGTCTCTCCCACACCCTCCATCTCT 5171
 445 ...............................  445
5170 CTCCCACGTGTCTCTCCCCTCACCTTCTCTCCCCTCCATTTCTCTCTCC 5121
 445 ...............................  445
5120 CTAATCTGTCTGTTCCCTGCCATGGCCCCCTTCTTCAAGCAGCCTCCC 5071
 445 ...............................  445
5070 ATCTTGCTCTCGCGGTCCCTCCCCGTCGTCTCTGTCTCACCCTGTTTCCA 5021
 445 ...............................  445
5020 CACCCTCACCTCTACCACCCCCTCAGCATGTTCCCTGGAAGCTGGAGGG 4971
 445 ...............................  445
4970 TCTCTGGGGCTCAGTCCCGGTCTCTCTTTCTCTCTCTCTCTCTCGTC 4921
 456 ..........CysValPheProArgGluLeuLysGluLysValPh 456
     |||||||||||||||||
4920 TCCCCGACCCTTCCCCCAGCGTGTTCCCGAGGAGCTGAAGGAGGGTGTT 4871
 456 eAlaSerTrpArgLeuArgCysAlaGlyArgGlyArgGlyAspIleAlaA 473
     ||||||||||||||||||||||||
4870 TGCTTCGTGGCCGGCTGCCGCGAGCCGGGAGGACATCGCAG 4821
 473 spArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIle 489
     ||||||||||||||||||||
4820 ACAGGCTTATCAGCGCCTCACTCTTCCTGCGCCTTCCTCTGCCCAGCGATT 4771
 490 MetSerProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnTh 506
     ||||||||||||||
4770 ATGTCGCCCAGTCTCTTTGGCTTATGCAGGAGTACCCAGATGAGCAGAC 4721
 506 rSerArgThrLeuThrLeuAlaLysValIleGlnAsnLeuAlaAsnP 523
     |||||||||||||||
4720 CTCACGAACCCTCACCCTCATTGCCAAGGTCATCCAGAACCTGGCCAACT 4671
 523 heSerLys 525
     |||||||
4670 TTTCCAAGTGAGGGAAGCTTCAGGAGGTGGGCAGGGAGTGGCCAGGG 4621
```

```
651  ............................................................  651
3720 TACAGTCAGTGGTAGGCTGAGAGCCCTTTACAGCCTGAGGGAGTGAGAGA          3671

651  ............................................................  651
3670 TTTGGAGCTCTAGGAATAGGGCTGAGGCTCCACCAACTCACGGCTTAGTT          3621

651  ............................................................  651
3620 GTAAGCCTAGAGCAGCATCCTGCTGCAAGCTCTGATTTGCTGTCCCTCTGCC       3571

651  ..........PheThrSerLysGluAspPheLeuGlyPheMetAs              537
3570 TGCCCATGCTAGTCCCCAGGCTGAGGTTCAGCCAGCCATGTCATGTCAGCC        3521

651  ............................................................  651
3520 ATGTGTCAAAATGTTCAAACATCTCAGTAATAGCTAGTGAATAAGCACTT         3471

651  ............................................................  651
3470 CCCCAGCCCCGACCACAACCCACAGACCTCCCATGATCCAGCACTT             3421

651  ............................................................  651
3420 AGAGCAGTGACAGCAGAAGCCTAGCAGGGCCTGCAGCCTGCTCCAGAGTC        3371

651  ............................................................  651
3370 CCAGCCTCCATTCTGATAGGTGGTGCCGTGCTTCATTTGCTGCTCATTA         3321

651  ............................................................  651
3320 TTTTGATGATGGCCCTGCTTTTCCTCTGGCCGTGCTTCCTCCATGACA         3271

651  ............................................................  651
3270 CCATACCCATCCCACCATTCCCGCCCTCTCCTTTCATTTGTCCACATCTCT      3221

652  .........................................SerSer1l         654
3220 CTCCTTCTCTGCTCTGTGCTCCGCCCCTCTTTCCATCTCTCTCCAGCTCCAT     3171

654  eAspLeuGlnSerPheMetAlaArgGlyLeuAsn..                        665
           |||||||||||||||||||||||||
3170 CGACCTTCAGTCCTTCATGGCTCGAGGGCTCAACAGGTGAGGGGCTCTCC       3121

665  ............................................................  665
3120 CCTCCCCGCCCTCCTCTCCTCTCCTCTCCCTCCTCCCCACTCCACT           3071

665  ............................................................  665
3070 GGCCTTCGCCCTACTCCTCCTCTCCTCCCTCCATGGACCTCATCTCCTC        3021

665  ............................................................  665
3020 CATATGTCCCAGCCCTGCCCCATCCTTCTCTTGCGCCCCATCTCC            2971

665  ............................................................  665
2970 CCTCCCTCTAGGCCTCACCCCTTCCCGGAGGGGGCCCTGTCCTTTCCCTTT      2921

665  ............................................................  665
2920 ACTCACCTGTCCCCTCCCATCCTCCTGCCCCTTCAGGGCTGCCA             2871

665  ............................................................  665
2870 CCGCTAGCTCTCCAGCCCTTCCCTCTCGGTCCCACTTTTCACCCAAGGCC       2821

665  ............................................................  665
```

```
525  ............................................................  525
4620 CAGGGAGTGGCGCAGGGCTCGGGGGTCGGCAAGAAGGGTCCTCCTGAGTCCCCA     4571

525  ............................................................  525
4570 GAGATCCTGAGATGGGGAGGCTATGATACCTTGTGTGTGTGTATGTGT          4521

525  ............................................................  525
4520 GTGTATGTGTGTGTGTGTGTGTGTGTGTGTATGGACCTTTATCTTCT           4471

526  ..........PheThrSerLysGluAspPheLeuGlyPheMetAs              537
           ||||||||||||||||||||||||
4470 GCATTCTTGGCTAGGTTTACCTCAAAGGAGGACTTTCTGGGCTTCATGAA         4421

537  nGluPheLeuGluLeuGluTrpGlySerMetGlnGlnPheLeuTyrGluI        554
           |||||||||||||||||||||||||||||
4420 TGAGTTTCTGGAGCTGGAATGGGGTTCCATGCACGCAGTTTTTGTATGAGA       4371

554  leSerAsnLeuAspThrLeuThrAsnSerSerPheGluGlyTyrIle          570
           |||||||||||||||||||||||||||||
4370 TCTCCAATCTGGACACGTTAACCAACAGTAGCTTTGAGGGTTACATC           4321

571  AspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLe        587
           ||||||||||||||||||||||||||||||
4320 GACTTGGGCCGAGAGCCTCTCCACACTGCCCTACTCTGGGAGGGGCT           4271

587  uProGlnLeuSerlys..Glu.                                     593
           |||||||||||     :::
4270 GCCCCAGCTCAGCAAGGTCAGCAGAGATCCCCTCTTTGCCCTATCCCCAGAT      4221

593  ............................................................  593
4220 GGCTCCAGAGGTTCCTGGAGACCTGAGAAAACTACCCTTTGAAGAATTTTTT      4171

593  ............................................................  593
4170 TCTCCCCTTGTTTCTCGAGGTGTCACCACTACTATCCCAACTCAGGGCCCC      4121

593  ............................................................  593
4120 CTCCACCTGCACCCTCAGAGACCCCTTCTTAGAGCTGGGCACTGGGAGCCCCCA    4071

594  ..........AlaLeuLeuLysLeuGlyProLeuP                        602
               |||||||||||||||||||
4070 GGTAACAGCCTCACCCTTCCAGGAAGCCCCTTCCTGAAGCTGGGTCCACTGC      4021

602  roArgLeuLeuSerAspIleSerThrAlaLeuArgAsnProAsnIleGln        618
           |||||||||::::||||||||||||||||||
4020 CCCGGCTCCTCAACGACATCAGCACAGCTCTGAGGAACCCCAACATCCAA        3971

619  ArgGlnProSerAlaArgGlnSerGluArgAlaArgSerGlnProMetValLe     635
           ||||||||||||||||||||||| :::||||||
3970 AGGCAGCCAAGCCCGCCAGAGTGAGCGGCCCCGGCCTCAGCCTGTGGTACT       3921

635  uArgGlyProSerArgAlaGluMetGlnGlyTyrMetMetArgAspLeuAsn.     651
           |||||||||||||||||||||||||||||||
3920 GCGGGGCCCATCGGCTGAGATGCAGGGCTACATGATGCGGGACCTCAACA        3871

651  ............................................................  651
3870 GGTGAGCAGCCTGGGACAGCCAGGCCTGCCCTGAGGCCCTTCCTCCTA          3821

651  ............................................................  651
3820 TTCTAGATACTCCTCACTGGGCCCCACCATGATCTCTTCTAGGGCTTGAAA        3771

651  ............................................................  651
3770 GAAGGGGAGGAAAAAGCACCAAGTTCTCAGGGGACGATAAGGAGACAGG         3721
```

us-09-294-298-6.rge

```
2820 TGTGCCAGACCACGACAAGGTTCAATTGCTAGGAGCCCTGACCTTACCTT 2771    1920 AGCGGTGGGGCAGCGGTGGGGGTGGGGGTGGGGGGCCAGCCGCCTCCATTGCA 1871
                                                               956 nArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysProArgP 973
 666 ..............Ser .SerMetAspMetAlaArgL 673          1870 GAGGGGCAAGTCTCAGCAGTTGACAGTCAGCGAGCCAGCCCAGAAACCCCGGC 1821
                  ::: |||||||||||||||
2770 CTGCTTGTGTGCCCCTTCCCCTTCTGACAGTCTATGGACATGGCTCGCC 2721   973 roSerSerGlyAsnLeuLeuGlnSerProGluProSerTyrGlyProAla 989
                                                               ||||||||||||||||||||||||||||||||||||||||||||||||||
 673 euProSerProThrLysGluLysProProProProGlyGlyGly 689          1820 CATCCAGCGGGAATCTATTGCAGTCCCCAGAGCCAAGTTATGGCCCCGCC 1771
     ||||||||||||||||||||||||||||||||||||||
2720 TCCCCTCCCAACCAAGGAAAAGCCACCCCCACCACGCCTGGTGTGGT 2671       990 ArgProArgGlnGlnSerLeuSerLysGluGlySerIleGlyLySerGl 1006
                                                               ||||||||||||||||||||||||||||||||||||||||||||||||||
 690 LysAspLeuPheTyrValSerArgProProLeuAlaArgSerProAl 706      1770 CGTCCACGGCAACAGAGCCTCAGCAAGGAGGGCAGCATTGGGGGCAGCGG 1721
     |||||||||||||||||||||||||||||||||||||||||||||||
2670 AAAGACCTGTTCTATGTAAGCCGTCCACCCCTGGCCCGTTCCTCACCAGC 2621    1006 yGlySerGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleThrLys 1022
                                                               ||||||||||||||||||||||||||||||||||||||||||||||||||
 706 aTyrCysThrSerSerSerAspIleThrGluProGlnLysMetLeuS 723      1720 GGGCAGCGCGGTGGCGGAGGGGGTGGGGGGCTGAAGCCCTCCATCACCAAGC 1671
     |||||||||||||||||||||||||||||||||||||||||||||||
2620 ATACTGCACGAGCAGCTCGGACATCACAGAGCCAGAGCAGAGATGCTGA 2571    1022 ..................................................  1022
                                                               
 723 erValAsnLysSerValSerMetLeuGlnGlyAspLeuGlyProGly 739      1670 AGGTAGGTGAAGGCAGGAGGAGGAGGCGGGCTGGGTCACAACAGGGAGGGAGGGAA 1621
     |||||||||||||||||||||||||||||||||||||||||||||||
2570 GTGTCAACAAGAGTGTGTCCATGCTGGACTTACAGGGTGATGGGCCTGGT 2521    1022 ..................................................  1022
                                                               
 739 GlyArgLeuAsnSerSerValSerAsnLeuAlaAlaValGlyAspLe 756      1620 GAAGGAGAATGGGGGGTGGGTTGAAACAGAGTCTGTGGCCTGAAGTTACA 1571
     |||||||||||||||||||||||||||||||||||||||||||||||
2520 GGCCGCTCAACAGCAGCAGTGTTTCGAACCTGGCCGGCCGTAGGGGACCT 2471    1022 ..................................................  1022
                                                               
 756 uLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProA 773   1570 ATCTTCTTGCCTCCTTTTGGCCATTAAACAAATGTGTATAGAGGGCCTGC 1521
     |||||||||||||||||||||||||||||||||||||||||||||||||
2470 GCTCACTCAAGCCAGGCCTCCCGACAGCAGCCTTGGGGCTACGGCCTG 2421     1022 ..................................................  1022
                                                               
 773 laProAlaGlyArgLeuSerGlnGlyGlySerSerIleThrAlaAla 789     1520 TATGTGCCATGTCTATGCCAGGCACTAAGGATACGGCACTGAACCCTCT 1471
     ||||||||||||||||||||||||||||||||||||||||||||||
2420 CGCCTGCCGACGCCTCCCAGGGAGTGGCTCATCCATCACCGGCGGCT 2371      1022 ..................................................  1022
                                                               
 790 GlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGl 806   1470 TAGCACTCTCATTCCAGAGGGGATTAATCCATAAGTAGAATGGGGTGAC 1421
     |||||||||||||||||||||||||||||||||||||||||||||||||
2370 GGCATGTCGCCTCAGCCAGATGGGTGTCACCACAGACGGTGTCCCTGCCCA 2321    1022 ..................................................  1022
                                                               
 806 nGlnLeuArgIleProLeuSerPheHisGlnAsnProLeuPheHisMetAlaA 823  1420 TGGAATAGGGCCAGCTGGGGCTTGGGAGGCATGAAGATCAGATCCTGAAT 1371
     |||||||||||||||||||||||||||||||||||||||||||||||||||
2320 GCAACTGGCGAATCCCCTCTCCTTCCAGAACCCTCTTCCACATGGCTG 2271     1022 ..................................................  1022
                                                               
 823 laAspGlyProGlyProAlaGlyHisHisGlyGlySerGlyHisGly 839     1370 CAGAACAACAGGAGAAATTCTCTGGCATTCAACTCACATCTCTGGCATTC 1321
     |||||||||||||||||||||||||||||||||||||||:::::||||||
2270 CTGATGGGGCCAGTGTCCCCAGGCGGCCATGGAGGGGGCGGTGGCCATGGC 2221    1022 ..................................................  1022
                                                               
 840 ProProSerSerHisHisHisHisHisHisHisHisHisHisArgGlyGl 856    1320 AGAGGTGATAAAGGCTACAGCAGGGCAATAGGACTGGGACTGTCGTCTCC 1271
     ||||||||||||||||||||||||||||||||||||||||||||||||||
2220 CCACCTTCCTTCCTCCCATCACCACCACCACCACCACCACCGAGGTGG 2171      1022 ..................................................  1022
                                                               
 856 yGluProGlyAspThrPheAlaProHeHisGlyTyrSerLysSerSerG 873    1270 TTTGGCGTGTCGTGTTGCCCTCTAAATGTACCCTGCTTTTTCCCACCCTTT 1221
     |||||||||||||||||||||||||||||||||||||||||||||||||
2170 AGAGCCCCCTGGGGACCACCTTTGCCCATTCCATGGCTATAGCAAGAGTG 2121    1023 ..............GlnHisSerGlnThrPro 1028
                                                               |||||||||||
 873 luAspLeuSerThrGlyValProLysProAlaAlaSerIleLeuHis 889     1220 CTTTCTCTGTTCGCCCTCACTGTGCCTTGCCATTCTCAGACACCA 1171
     ||||||||||||||||::||||||||||||||||||||||||||||
2120 AGGAACCTCTCTTCCGGGGTCCCCAGGCGGTCCCCAAGCCCCTGCTGCCCTTCAT 2071    1029 SerThrLeuAsnProThrMetProAlaSerGluArgThrValAlaTrpVa 1045
                                                               ||||||||||||||||||||||||||||||||||||||||||||||||||
 890 SerHisSerTyrSerAspGluPheGlyThrAspPheThrAr 906           1170 TCCACATTGAACCCCACACATGCCAGCCCTCTGAGCGGACAGTGGCCTGGGT 1121
     ||||||||||||||||||||||||||||||||||||||||
2070 AGCCACAGCTACACGTGATGAGTTTGGACCCCTCTGGCACTGACTTCACCCG 2021    1045 lSerAsnMetProHisLeuSerAlaAspIleHisIleGluA 1062
                                                               ||||||||||||||||||||||||||||||||||||||||
 906 gArgGlnLeuSerLeuGlnAspAsnLeuHisMetLeuSerProProG 923     1120 CTCCAACATGCCTCACCTGTCGGCTGACATCCAGAGTGCCCACATGCGAGC 1071
     |||||||||||||||||||||||||||||||||||||||||||||||
2020 TCGGCAGCTTTCACTCACCAGGACAACCTGCCAGCACATGCGTCCCCCTCCCC 1971    1062 rgGluGluTyrLysLeuLysGluTyrSerLysSerMetAspGluSerArg 1078
                                                               ||||||||||||||||||||||||||||||||||||||||||||||||||
 923 lnIleThrIleGlyProGlnArgProSerGlyProGlyGlyGlyGly 939     1070 GGGAAGAGTACAAGCTCAAGGAGTACTCAAAATCGATGGATGAGAGCCGG 1021
     |||||||||||||||||||||||||||||||||||||||||||||||
1970 AGATCACCATTGGTCCCCAGAGGGCCAGCCCCCTCAGGGCCTGGAGGTGGG 1921    1079 LeuAspArgValLys 1083
                                                               |||||||||:::
 940 SerGlyGlyLySerGlyGlyGlyGlyGlyGlyGlyLyGlyGlnProProLeuGl 956  1020 CTGGATAGGGTACGA 1006
     ||||||||||||||||||||||||||||||||||||
```

```
seq_name: gb_htg20:AL161906

seq_documentation_block:
LOCUS       AL161906   169323 bp   DNA           HTG        18-AUG-2000
DEFINITION  Homo sapiens chromosome 6 clone RP11-567N9, *** SEQUENCING IN
            PROGRESS ***, 24 unordered pieces.
ACCESSION   AL161906
VERSION     AL161906.5  GI:9863619
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 169323)
AUTHORS     Sims,S.
TITLE       Direct Submission
JOURNAL     Submitted (17-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
            On Aug 21, 2000 this sequence version replaced gi:9212890.
COMMENT     Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ------------- Project Information
            Center project name: bA567N9
            ------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: plasmid; L08752; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Consensus quality: 155424 bases at least Q40
            Consensus quality: 160003 bases at least Q30
            Consensus quality: 163252 bases at least Q20
            Insert size: 167023; sum-of-contigs
            Insert size: 186175; 5.7% error; agarose-fp
            Quality coverage: 3.59x in Q20 bases; sum-of-contigs Quality
            coverage: 3.22x in Q20 bases; agarose-fp
*
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1    5456: contig of 5456 bp in length
*     5457    5556: gap of     100 bp
*     5557    8767: contig of 3211 bp in length
*     8768    8867: gap of     100 bp
*     8868   17168: contig of 8301 bp in length
*    17169   17268: gap of     100 bp
*    17269   26591: contig of 9323 bp in length
*    26592   26691: gap of     100 bp
*    26692   31188: contig of 4497 bp in length
*    31189   31288: gap of     100 bp
*    31289   37579: contig of 6291 bp in length
*    37580   37679: gap of     100 bp
*    37680   45350: contig of 7671 bp in length
*    45351   45450: gap of     100 bp
*    45451   48354: contig of 2904 bp in length
*    48355   48454: gap of     100 bp
*    48455   63401: contig of 14947 bp in length
*    63402   63501: gap of     100 bp
*    63502   67724: contig of 4223 bp in length
*    67725   67824: gap of     100 bp
*    67825   70550: contig of 2726 bp in length
*    70551   70650: gap of     100 bp
*    70651   80664: contig of 10014 bp in length
*    80665   80764: gap of     100 bp
*    80765   83307: contig of 2543 bp in length
*    83308   83407: gap of     100 bp
*    83408   88383: contig of 4976 bp in length
*    88384   88483: gap of     100 bp
*    88484   93697: contig of 5214 bp in length
*    93698   93797: gap of     100 bp
*    93798   99354: contig of 5557 bp in length
*    99355   99454: gap of     100 bp
*    99455  111055: contig of 11601 bp in length
*   111056  111155: gap of     100 bp
*   111156  118645: contig of 7490 bp in length
*   118646  118745: gap of     100 bp
*   118746  122421: contig of 3676 bp in length
*   122422  122521: gap of     100 bp
*   122522  132468: contig of 9947 bp in length
*   132469  132568: gap of     100 bp
*   132569  137959: contig of 5391 bp in length
*   137960  138059: gap of     100 bp
*   138060  151691: contig of 13632 bp in length
*   151692  151791: gap of     100 bp
*   151792  158253: contig of 6462 bp in length
*   158254  158353: gap of     100 bp
*   158354  169323: contig of 10970 bp in length.

FEATURES             Location/Qualifiers
     source          1..169323
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-567N9"
                     /clone_lib="RPCI-11.2"
     misc_feature    1..5456
                     /note="assembly_fragment:00884
                     fragment_chain:1"
     misc_feature    5557..8767
                     /note="assembly_fragment:00832
                     fragment_chain:1"
     misc_feature    8868..17168
                     /note="assembly_fragment:01060
                     fragment_chain:1"
     misc_feature    17269..26591
                     /note="assembly_fragment:01172
                     fragment_chain:1"
     misc_feature    26692..31188
                     /note="assembly_fragment:00279
                     fragment_chain:1"
     misc_feature    31289..37579
                     /note="assembly_fragment:01205
                     fragment_chain:1"
     misc_feature    37680..45350
                     /note="assembly_fragment:01334
                     fragment_chain:2"
     misc_feature    45451..48354
                     /note="assembly_fragment:01263
                     fragment_chain:2"
     misc_feature    48455..63401
                     /note="assembly_fragment:00990
                     fragment_chain:2"
     misc_feature    63502..67724
                     /note="assembly_fragment:00045
                     fragment_chain:2"
     misc_feature    67825..70550
                     /note="assembly_fragment:00965
                     fragment_chain:2"
     misc_feature    70651..80664
                     /note="assembly_fragment:00637
                     fragment_chain:3"
     misc_feature    80765..83307
                     /note="assembly_fragment:01291
                     fragment_chain:3"
     misc_feature    83408..88383
                     /note="assembly_fragment:01317
                     fragment_chain:3"
     misc_feature    88484..93697
                     /note="assembly_fragment:01171
```

us-09-294-298-6.rge

Fri Jan 19 06:20:05 2001

```
       misc_feature    fragment_chain:4"
                        93798..99354
       misc_feature    /note="assembly_fragment:00288
                        fragment_chain:4"
       misc_feature    99455..111055
                        /note="assembly_fragment:00286
                        fragment_chain:4"
       misc_feature    11156..118645
                        /note="assembly_fragment:00303
                        fragment_chain:5"
       misc_feature    118746..122421
                        /note="assembly_fragment:01184
                        fragment_chain:5"
       misc_feature    122522..132468
                        /note="assembly_fragment:00501
                        fragment_chain:6"
       misc_feature    132569..137959
                        /note="assembly_fragment:01244
                        fragment_chain:6"
       misc_feature    138060..151691
                        /note="assembly_fragment:00032"
       misc_feature    151792..158253
                        /note="assembly_fragment:00779"
       misc_feature    158354..169323
                        /note="assembly_fragment:01270"
BASE COUNT    41674 a   41332 c   41269 g   42673 t   2375 others
ORIGIN

alignment_scores:
       Quality:      3987.00          Length:   2457
       Ratio:        4.237            Gaps:     12
Percent Similarity:  38.299    Percent Identity:  37.810

alignment_block:
US-09-294-298-6 x AL161906     ..
Align seg 1/1  to: AL161906  from: 1  to: 169323
```

```
  139 ProAsnLysAspAsnSerArgArgValAspAsnValLeuLysLeuTrpIl 155
      ||| :::|||||||||||||||||||||||||||||||||||||||||
77990 CCATGCCAGCAACAGCCCGCGACAACAATGTGCTAAAGCTGTGGGAT 78039

  155 eIleGluAlaArgGluLeuProProLysLysArgTyrTyrCysGluLeuC 172
      |||||||||||||||||||||||||||||||||||||||||||||||||
78040 CATAGAGGCCCGGAGCTGCCCCCCAAGAAGCGGTACTACTGTGAGCTCT 78089

  172 ysLeuAspAspMetLeuTyrArgAlaArgThrThrSerLysProArgSerAla 188
      |||||||||||||||||||||||||||||||||||||||||||||||||
78090 GCCTGGATGACAATGCTGTATGCACGCACCACCTCCAAGCCCCGCTCTGCC 78139

  189 SerGlyAspThrValPheTrpGlyGluHisPheGluPheAsnAsnLeuPr 205
      |||||||||||||||||||||||||||||||||||||||||||||||||
78140 TCTGGGGACACCGTCTTCTGGGGCGAGCACTTCGAGTTTAACAACCTGCC 78189

  205 oAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysA 222
      |||||||||||||||||||||||||||||||||||||||||||||||||
78190 GGCTGTCCGTGCCCTGCGCCTGCATCTGTACCGTGACTCAGACAAAAAGC 78239

  222 rgLysLysAspLysAlaGlyTyrValGlyLeuValThrValProValAla 238
      |||||||||||||||||||||||||||||||||||||||||||||||||
78240 GCAAGAAGGACAAGGCAGGCTATGTCGGCCTGGTGACGGTGCCAGTGGCC 78289

  239 ThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeuPr 255
      |||||||||||||||||||||||||||||||||||||||||||||||||
78290 ACCCTGGCTGGCCGGCACTTCACAGAGCAGTGGTACCCTGTAACCCTGCC 78339

  255 oThrGlySerGlyGlySerGlyGlyMetGlySerGlyGlyGlyGlyGlyS 272
      |||||||||||||||||||||||||||||||||||||||||||||||||
78340 AACAGGCGGATCTGGGGGCCATGGGTTCGGAGGGGAGGGGAGGGGGGCT 78389

  272 erGlyGlySerGlyGlyLysGlyLysGlyGlyGlyCysProAlaValArg 288
      |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
78390 CGGGGGTGGCTCAGGGGCAAGGGCAAAGGAGGGAGGTTGCCGGCGTGTGCGG 78439

  289 LeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLy 305
      |||||||||||||||||||||||||||||||||||||||||||||||||
78440 CTGAAAGCACGTTACCAGACAATGAGCATCTTGCCCATGGAGCTATATAA 78489

  305 sGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaValL 322
      |||||||||||||||||||||||||||||||||||||||||||||||||
78490 AGAGTTTGCAGAGTATGTCACCAACCATTATCGGATGCTGTGTGCAGTCT 78539

  322 euGluProAlaLeuAsnValLysGlyLysGluGluValAlaSerAlaLeu 338
      |||||||||||||||||||||||||||||||||||||||||||||||||
78540 TGGAGCCCGCCCTGAATGTCAAGGGCAAGGAGGAGGTTGCCAGTGCCACTA 78589

  339 ValHisIleLeuGlnSerThrGlyLysAla.Lys............... 349
      |||||||||||||||||  :::
78590 GTTCACATCCTGCAGAGTACAGGCCAAGGTGAGTGTTGTGCCCTC 78639

  349 ................................................. 349
78640 AGGGAAAGGTGACTTGGGAATGGCACTTGCTTGGGGGTTAGTGAGGACA 78689

  349 ................................................. 349
78690 GGGCAAATTCACCGAGATTGGGTTGTGCAGAGGCTGACACTTGGATTTCC 78739

  350 ..........AspPheLeuSerAspMetAlaMetSerGluValAspArgP 363
                |||||||||||||||||||||||||
78740 TGGGCCTCAGGACTTCCTTTCAGACATGGCCATGTCTGAGGTAGACCGGT 78789

  363 heMetGluArgGluHisIleLeuIlePheArgGluAsnThrLeuAlaThrLys 379
      |||||||||||||||||||||||||||||||||||||||||||||||||
78790 TCATGGAACGGGAGCACATCCTCATATTCCCGCAGAACACGCCTTGCCCACTAAA 78839

  380 AlaIleGluGluTyrMetArgLeuIleGlyGlnLysLysTyrLeuLysAspAl 396
      |||||||||||||||||||||||||||||||||||||||||||||||||
78840 GCCATAGAAGAGTATATGAGACTGATTGGTCAGAAAATACCTCAAGGATGC 78889

  396 a.Ile............................................ 397
      | :::
78890 CATTGGTATGGCCACACTCAGGCCCTTCTTCCCCAAACCTGCCCAGATG 78939

  397 ................................................. 397
78940 TCCACCCCAGACCCCAAGTCACCCTTCCACAGCTTGTATCTTCCTAACC 78989

  397 ................................................. 397
78990 CAGAGTCCTAGGACTCCAGCCTCCAACACCTGACTTCTGAAATTCCCCAA 79039

  397 ................................................. 397
79040 CCCTGCCGCACCCCCTTCCCGCCCTTGGAAAGTGTGACCACACCCTCTTG 79089

  398 .............GlyGluPheIleArgAlaLeuTyrGluSerGluGl 409
                   |||||||||||||||||||||||||
79090 TGCCCCCACCCCCAGGAGAATTCATCCGTGCTCTGTATGAANTCTGAGGA 79139

  409 uAsnCysGluValAspProIleLysCysThrAlaSerLeuAlaGluH 426
      |||||||||||||||||||||||||||||||||||||||||||||||||
79140 AAACTGCGAGGTAGACCCTATCAAGTGCACAGCATCCAGTTTGGCAGAGC 79189

  426 isGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysValVal 442
      |||||||||||||||||||||||||||||||||||||||||||||||||
79190 ACCAGGCCAACCTGCGAATGTGTTGTGAGTTGGCCCTGCCAAGGTGGTC 79239

  443 AsnSerHis........................................ 445
      |||||||||
79240 AACTCCCACTGGTGAGACTGGGAACACGCTGGGCTGGGGGGGTCGGG 79289

  445 ................................................. 445
79290 GGAATTATGTTGTTCATCTGTTCATCTATCTGTCCATCCTCAAAGAGGACT 79339
```

```
445  ·······································
79340 GAGCACCATTTATGGGCAAAGCATTGTTCTAGGCGCTATAGAGCAAACAG 79389
445  ·······································
79390 GTGAAAGAGGCCTGGTCCCCTGCCCTCAGAGGGCCTCCACCAGAATGGGGA 79439
445  ·······································
79440 CAAATTAGAAGAAAAAAAAAAAGCCACAGAGCCATAATGTGTCTGTGTAAG 79489
445  ·······································
79490 TGCTGAGTAAGGGTCCCCCAACCTCTGTGTGACATAAGGTCAGAGAGAA 79539
445  ·······································
79540 GGCAGAGCTTTGAGATAAGTGGGGAAGAGGTGCCCCCTTGGGTAGGCTTT 79589
445  ·······································
79590 GAAGACTGGTTGGTTTAGGTTCTGATATATGGCAGATAGTTGGCAAAGACACAT 79639
445  ·······································
79640 TTCAGAAGAAGGCTGTGAGAAAGGCACATGTGTGATGGTGAAAAGGCCCA 79689
445  ·······································
79690 GGAGTTTTCAGGGGACACATTAAAGTAGGTTAGTAGCAATTACATCAGGTTT 79739
445  ·······································
79740 AGTGGAGCATGTGCCTCATAATGGGGAGCTGGCGGGAGAGATCTCTGGGCA 79789
445  ·······································
79790 GGAAGATTAGCTTTAGAAACTGGAAGGCCTCAAGGAGTCTGAGGTCATTG 79839
445  ·······································
79840 TTAGGCCTTGGGATGCCATTAAAGGTGTCAGAGAAATATTGTGATTTAGAAG 79889
445  ·······································
79890 ATTAATCCATAGGCTGGGCACCGGTGGCTCACACCTGTAATACCAGCACTT 79939
445  ·······································
79940 TGGGAGGCTGAGGCGGCGGGCAGACCACCTGAGGTCAGGAGTTTGAGACCAGG 79989
445  ·······································
79990 CTGACCAACATGGAGAAACCCTGTCTACTAAAAATACAAAAATTAGCCA 80039
445  ·······································
80040 GATGTGGTGGCACATGCCTGTAATCCCAGCTATTCGGGAGGCTAAGGCAG 80089
445  ·······································
80090 GAGAATCACTTGAATCTGGGAGGTGGAGGTTGTAGTGAGCCGAGATCACA 80139
445  ·······································
80140 CCATTGCACTCTAGCCTGGGCAACAAGAGTGAAACTTCCGTCTCAAAAAAAA 80189
445  ·······································
80190 AAAAATTAATCCATAAGTAGAGTGTATGTAGTAGAGGGGATCGTATG 80239
445  ·······································

445  ·······································
80240 TTGGGGGATGACTGGAATAGAGCCAGCTGGGGGCTCGAGGCATGAAAGT 80289
445  ·······································
80290 CAGATCCTGAATCAGAACAGTGACAGAAGTTAGGAAGGAGCTACTGGAAG 80339
445  ·······································
80340 GACCCTCGGGAAAGGATCAGTAGGAATTGTCAGCTTATTGGAAAGGAG 80389
445  ·······································
80390 GGAGCATGTCTCTGGGAGTCACGGATGACTCAAGAGGCCATGAGGGCTGG 80439
445  ·······································
80440 TGGTTGGGAGACCGGTGGCCTTCATTGACAACCAGGAAAGTCAGGAGGAGG 80489
445  ·······································
80490 AGCCAGTTGGAGGTGTGGGGGCGATGGTGAGCTCTTGCTTTACCACCAGCCG 80539
445  ·······································
80540 AGTTTAAGGTGTCAGTGGGACATTGAAGTGAACTGGGAGGTGCGGGGAG 80589
445  ·······································
80590 TGAGCTCTGAGCCATTCCAGGGACTGGGGATCATGCCTGGGGCACCTCCA 80639
445  ·······································
80640 TCCCCATTTCCCTGGATCCAGAAGANNNNNNNNNNNNNNNNNNNNNNNNN 80689
445  ·······································
80690 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 80739
445  ·······································
80740 NNNNNNNNNNNNNNNNNNNNNNNTCCCCTGNNATCCANAAAGTTGGGG 80789
445  ·······································
80790 GGTCCGAGCTCCCTGTACCTCAAGTGACCCTCCATCTCTCTCCCATCTCT 80839
445  ·······································
80840 GTCTCTCCCTGTGTGTCTGTTTTTTCTCTCCTCCTTGTCTCTCTCTCC 80889
445  ·······································
80890 CACACCCCTCCATCTCTCTCTCCTAATCTGTCTGTTCCCTCTCACCTTCTCTCC 80939
445  ·······································
80940 CCTCCATTTCTCTCTCCTAATCTGTCTGTTCCCTCTGCCATGGCCCCCT 80989
445  ·······································
80990 TCTTCAAGCAGCCTCCCATCTTGCTCTGCGCGTCCCTCCTTCCCGTCTC 81039
445  ·······································
81040 TCTCACCCCTGTTCCACACCCTCACCTCCTACCACCCCCCAGCATGT 81089
445  ·······································
81090 TCCCTGGAAGCTGAGGCTCTCTGGGGCTCAGTCCCGGTCTCTCTTTCT 81139
446  ·······································  CysValPheProArg 450
```

```
                                       |||||||||
81140 CTCTCTCTCTCTCTGTCTCCCCGACCCTTCCCCCAGCGTGTTCCCGAGG  81189

451   GluLeuLysGluValPheAlaSerTrpArgLeuArgCysAlaGluArgGl  467
      ||||||||||||||||||||||||||||||||||||||||||| |||||
81190 GAGCTGAAGGAGGTGTTTGCTTCATGGCGGCTGCGCTGCGCAGAGCGAGG  81239

467   yArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuArgP  484
      ||||||||||||||||||||||||||||||||||||||||||||||||||
81240 CCGGGAGGACATCGCAGACAGGCTTATCAGCGCCTTCACTCTTTCCTGCCT  81289

484   heLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGlnGlu  500
      ||||||||||||||||||||||||||||||||||||||||||||||||
81290 TCCTCTGCCCAGCCGATTATGTCGCCCCAGTCTCTTTGGGCTTATGCAGGAG  81339

501   TyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysValIl  517
      ||||||||||||||||||||||||||||||||||||||||||||||||
81340 TACCCAGATGAGCAGACCAGCCTCACGAACCCTCACCCTCATTGCCAAGGTCAT  81389

517   eGlnAsnLeuAlaAsnPheSerLys.............               525
      ||||||||||||||
81390 CCAGAACCTGGCCAACTTTCCAAGTGAGGAAGCTTCAGGAGTGGGCAG    81439

525   .................................                   525

81440 GGCAGGAGTGGCCAGGGGAGGGAGTGGCCAGGGGTCGGGCAAGAAG      81489

525   .................................                   525

81490 GGTCTCCTGAGTCCCCAGAGAATCCTGAGATGGGAGGCTATGATACCTTG  81539

525   ....................................               525

81540 TGTGTGTGTGTGTGTATGTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTA  81589

526            PheThrSerLysGluAsp                       531
                ||||||||||||
81590 TGTGACCTTTATCTTCTGCCATTCTTGGCTAGGTTTACCTCAAAGGAGGAC  81639

532   PheLeuGlyPheMetAsnGluPheLeuGluLeuGluTrpGlySerMetGl  548
      ||||||||||||||||||||||||||||||||||||||||||||||||
81640 TTTCTGGGCTTCATGAATGAGTTTCTGGAGCTGGAATGGGGTTCCATGCA  81689

548   nGlnPheLeuTyrGluIleSerAsnLeuAspThrLeuThrAsnSerSerS  565
      ||||||||||||||||||||||||||||||||||||||||||||||||
81690 GCAGTTTTTGTATGAGATCTCCAATCTGGACACGCTAACCAACAGCAGTA  81739

565   erPheGluGlyTyrIleAspLeuGlyArgGlyLeuGluSerThrLeuHisAla  581
      ||||||||||||||||||||||||||||||||||||||||||||||||
81740 GCTTTGAGGGTTACATCGACTTGGGCCGAGAGCTCTCCACACTGCATGCC  81789

582   LeuLeuTrpGluValProGlnLeuSerLys..Glu..............   593
      ||||||||||||||||||||||||||| |||
81790 CTACTCTGGGAGGTGCTGCCCCAGCTCAGCAAGGTCAGCAGCAATCCCCTCT  81839

593   ...............................                     593

81840 TTGCCCTATCCCAGATGGCTCCAGAGGTTCCTGGAGCCTGAGAAACTAC  81889

593   ...............................                     593

81890 CCTTTGAAGATTTTTTTTCTCCCTGTTTCTCGAGGTGTCACCACTACT  81939

593   ...............................                     593

81940 ATCCCAACTCAGGCCCCCTTCCACCCTGCACCCTCTTAGAGC          81989

594                            AlaLeuLe                  596
                              ||||||
81990 TGGGCACTGAGCCCCCAGGTAACAGCCCTCACCCTTCCAGGAAGCCCTCCT  82039

596   uLysLeuGlyLeuProArgLeuLeuSerAspIleSerAspIleSerThrAlaLeuA  613
      ||||||||||||||||||||||||||||:::||||||||||||||||||
```

```
                                         |||||||||
82040 GAAGCTGGGTCCACTGCCCGGCTCCTCAACGACATCAGCACAGCTCTGA  82089

629   rgAsnProAsnIleGlnArgGlnSerGluArgSerGluArgAlaArg     629
      ||||||||||||||||||||||||||||||||||||||||       |||
82090 GGAACCCCAACATCCAAAGCAGCCAAGCCCCAGAGTGAGGCGGCCCGG  82139

630   SerGlnProMetValLeuArgGlyProSerAlaGluMetGlnGlyTyrMe  646
      ||||||:::||||||||||||||||||||||||||||||||||||||||
82140 CCTCAGCGTGGTACTGCGGGGGCCATCGGCTCGAGATCAGGGCTACAT  82189

646   tMetArgAspLeuAsn...............                     651
      ||||||||||||||
82190 GATGCGGGACCTCAACAGGTGAGCACCCTGGGACAGCCAGGCCTGTGCCC  82239

651   ....................................               651

82240 TAGGAGCCCTTCTCCTATTCTTAGTACTCCTCACTGGGCCCCACATGCAT  82289

651   ....................................               651

82290 CTCTCTAGGGCTTGAAAGAGGAGGAGGAAAAAGCACCAAGTTCTCAGGGGA  82339

651   ....................................               651

82340 GACGATAAGGAGGACAGGTACAGTCAGTGGTGAGGCTGAGAGCCCTTTACAG  82389

651   ....................................               651

82390 CCTGAGGGAGTGAGAGATTTGGAGCTCTAGGAATAGGGCTGAGGCTCCAC  82439

651   ....................................               651

82440 CAACTCACGGCTTAGTTGTAAGCCTAGAGCATCCCTGCTGCAAGCTCTGA  82489

651   ....................................               651

82490 TTTGCTGTCCCCTGCCCATGCTAGTCCCCAGGCTGAGGTTCAGCC      82539

651   ....................................               651

82540 AGCATGTCATGTCAGCCATGTGTCAAAATGTTCAAACATCTCAGTAATAG  82589

651   ....................................               651

82590 CTAGTGAATAAGCACTTCCCCCAGCCCCCGACCACAACCCCACAGACCTC  82639

651   ....................................               651

82640 CCCATGATCCAGCACTTAGAGCACTTGACAGCAGAAGCCTAGCAGGGCCTG  82689

651   ....................................               651

82690 CAGCCCTGCTCCCAGAGTCCCAGCCTCCATTCTGATAGGTGGTGCCCGTGCT  82739

651   ....................................               651

82740 TCATTTGCTGCTCATTATTTTGATGATGGCCCTGCTTTTTCCTGTCCCCG  82789

651   ....................................               651

82790 TGTCTTCCTCCATGACGACCATTACCCACCATTCCCGCCTCCTCCTCTT  82839

651   ....................................               651

82840 CATTTGTCCACATTCTCTCTCTCTCTCTGTCTGTCGCCCCCTCTTTCCA  82889

652   ........SerSerIleAspLeuGlnSerPheMetAlaArgGlyLeuAs   665
                      ||||||||||||||||||||
82890 TCTCTCTCCCAGCTCCATTCCAGCTTCCAGTCCTTCCAGTCCTTCATGGCTCGAGGCCTCAA  82939

665   n.............                                      665
      -
82940 CAGGTGAGGGGGCTCTCCCCTCCCCCCGCCTCCTCCTCCTCCTCGTCTGTT  82989
```

Fri Jan 19 06:20:05 2001

us-09-294-298-6.rge

```
756  uLeuHisSerSerGlnAlaSerLeuThrAlaAlaAlaGlyLeuArgProA  773
     |||||||||||||||||||||||||||||||||||||||||||||||||
83890 GCTGCACTCAAGCCAGCCAGGCCTGCTGACAGCAGCCTGGGGCTACGGCCTG 83939

773  laProAlaGlyAlaArgLeuSerGlnGlySerGlySerIleThrAlaAla  789
     |||||||||||||||||||||||||||||||||||||||||||||||||
83940 CGCCTGCCGGACGCGCTCTCCCAGGGGAGTGGCTCATCCATCACGGCGGCT 83989

790  GlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGl  806
     |||||||||||||||||||||||||||||||||||||||||||||||||
83990 GGCATGCGCCTCAGCCAGATGGGTGTCACCACAGACGGTGTCCCTGCCCA 84039

806  nGlnLeuArgIleProLeuSerProPheHisMetAlaA  823
     |||||||||||||||||||||||||||||||||||||
84040 GCAACTGCGAATCCCCCTCTCCTTCCAGAACCCTCTTCCACATGGCTG 84089

823  laAspGlyProGlyProAlaGlyHisGlyGlySerGlyHisGly  839
     ||||||||||||||::::::::||||||||
84090 CTGATGGGCCCAGGTCCCCAGGCGGCCATGGAGGGGCGGTGGCCATGGC 84139

840  ProProSerHisHisHisHisHisHisHisHisHisArgGlyGl  856
     |||||||||||||||||||||||||||||||||||||||||||
84140 CCACCTTCCTCCCATCACCACCACCACCACCACCACCACCGAGGTGG 84189

856  yGluProProGlyAspThrPheAlaProPheHisGlyTyrSerLysSerG  873
     |||||||||||||||||||||||||||||||||||||||||||||||||
84190 AGAGCCCCTGGGACACCTTTGCCCCATTCCATGGCTATAGCAAGAGTG 84239

873  luAspLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHis  889
     ||||||||||||:::::||||||||||||||||||||||||||||||
84240 AGGACCTCTCTTCCGGGGTCCCCAAGCCCCCTGCCGCCTCCATCCTTCAT 84289

890  SerHisSerTyrSerAspGluPheGlyProSerGlyThrAspPheThrAr  906
     |||||||||||||||||||||||||||||||||||||||||||||||||
84290 AGCCACAGCTACAGTGATGAGTTTGGACCCTCTGGCACTGACTTCACCCG 84339

906  gArgGlnLeuSerLeuGlnAspAsnGlnHisMetLeuSerProProG  923
     |||||||||||||||||||||||||||||||||||||||||||||
84340 TCGGCAGCTTTCACTCCAGGACAACCTGCACACATGCTGTCTCCCTCCCC 84389

923  lnIleThrIleGlyProGlnArgProAlaProSerGlyGlyGlyGlyGly  939
     |||||||||||||||||||||||||||||||||||||||||||||||||
84390 AGATCACCATTGGTCCCCAGAGGCCAGCCCCCTCAGGGCCTGAGGTGGG 84439

940  SerGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProLeuGl  956
     |||||||||||||||||||||||||||||||||||||||||||
84440 AGCGGTGGGGCAGCCGGTGGGGGTGGCGGGGGCCAGCCGCCTCCATTGCA 84489

956  nArgGlyLysSerGlnLeuThrValSerAlaAlaGlnLysProArgP  973
     |||||||||||||||||||||||||||||||||||||||||||||
84490 GAGGGCCAAGTCTCAGCAGTTGACAGTGGCAGCCCAGAAACCCCGGC 84539

973  roSerGlyAsnLeuLeuGlnSerProGluProSerTyrGlyProAla  989
     ||||||||||||||||||||||||||||||||||||||||||||||
84540 CATCCAGCGGGAATCTATTGCCAGTCCCCAGAGCCAAGTTATGGCCCCGC 84589

990  ArgProArgGlnGlnSerLeuLysSerGluGlySerIleGlyGlySerGl  1006
     |||||||||||||||||||||||||||||||||||||||||||||||||
84590 CGTCCCACGGCAACACAGACCTCAGCAAGGAGGGCAGCATTGGGGACGCGG 84639

1006 yGlySerGlyGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleThrLys.  1022
     ||||||||||||||||||||||||||||||||||||||||||||||||
84640 GGGCAGCCGGTGGCGAGGGGGTGGGGGGTGAAGCCCTCCATCACCAAGC 84689

1022 ......  1022
84690 AGGTAGGTGAAGGCAGGAGGAAGGCGGGCTGGGTCACACAGGAGGGGAA 84739

1022 ......  1022
84740 GAAGGAGGATGGGGGTGGGGTTGAAACAGAGTCTGTGGCCTGAAGTTACA 84789

1022 ......  1022
```

```
665  ......  665
82990 CCCTCTCCCACTCCACTGGCCTTCGCCCCTACTCCTCTCCCTCCTTC 83039

665  ......  665
83040 ATGGACCTCATCTCCTCCATATGTGCCCAGCGCCTGCCCCCATCCCTTCTG 83089

665  ......  665
83090 TTGCTGCCCCCATCTTCCCTCCTCCTAGGGCTTACCCCCTTTCCGGAGGGG 83139

665  ......  665
83140 CCCTGTCCTTTTCCCTTTACCTTACCTGTCCCCTCCCCTTCCTCCCTGCCTG 83189

665  ......  665
83190 CTCTTTTTAGGCTGTGTCACCGGTTAGCTTCAGCCCTCCCTCTGGTCCCC 83239

665  ......  665
83240 TTTTTACCCCAGGGCCTTGTGTGCCACCACAACAAGGGTCAATTGGTAGGA 83289

665  ......  665
83290 GCCCTTACCTAACCTTTGNNNNNNNNNNNNNNNNNNNNNNNNNNN 83339

665  ......  665
83340 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 83389

665  ......  665
83390 NNNNNNNNNNNNNNNCCTTCCCGAAGGGGCCCTGTCCTTTCCCTTT 83439

665  ......  665
83440 ACTCACCTGTCCCCTCCCATCCTCCCTGCCCTCCTTCAGGGCTGCCA 83489

665  ......  665
83490 CCGCTAGCTCTCAGCCTCCCTGGGTCCCACTTTTCACCCCAAGGCC 83539

665  ......  665
83540 TGTGCCAGACCACAGCAAGGTTCAATTGCTAGGAGCCCTTGACCTTACCTT 83589

666  Ser..SerMetAspMetAlaArgL  673
          |||||||||||||||
          :::
83590 CTGCTGTGTGCCCCTTCCCCTTCTGACAGCTCTATGGACATGGCTCGCC 83639

673  euProSerProThrLysGluProProProGlyProGlyGly  689
     |||||||||||||||||||||||||||||||||||||||
83640 TCCCCCTCCCCACCAACAGGAAAGCCACCCCCACCACCCGCTGGTGGTGGT 83689

690  LysAspLeuPheTyrValSerArgProProLeuAlaArgSerProAl  706
     ||||||||||||||||||||||||||||||||||||||||||||||
83690 AAAGACCTGTTCTATGTAAGCCCGTTCCACCCCTGGCCCGTTCCTCACCAGC 83739

706  aTyrCysThrSerSerSerIleGluProGluGlnLysMetLeuS  723
     ||||||||||||||||||||||||||||||||||||||||||
83740 ATACTGCACAGCAGCTCGGACATCACAGAGCCAGAAGATGCTGA 83789

723  erValAsnLysSerValSerMetLeuAspProGluGlnGlyAspGlyProGly  739
     |||||||||||||||||||||||||||||||||||||
83790 GTGTCAACAAGAGTGTGTCCATGCTGGACCCTGAGCAGGGTGGGCCTGGT 83839

740  GlyArgLeuAsnSerSerValSerAsnLeuAlaValAlaGlyAspLe  756
     ||||||||||||||||||||||||||||||||||||||||||||||
83840 GGCCGCCTCAACACAGCAGTGTTTCGAACCTGGCCGGCCGTAGGGACCT 83889
```

```
84790  ATCTTCTTGCCTCCTTTGGCCATTAAACAAATGTGTATAGAGGGGCCTGC  84839
1022   .................................................  1022
84840  TATGTGCCATGTGCTATGCCAGGCACTAAGGATACGGCACTGAACCCTCT  84889
1022   .................................................  1022
84890  TAGCACTCTCATTCCAGAGGGATTAATCCATAAGTAGAATGGGGGTGAC   84939
1022   .................................................  1022
84940  TGGAATAGGGCCAGCTGGGGGCTTGGAGGCATGAAGATCAGATCCTGAAT  84989
1022   .................................................  1022
84990  CAGAACAACAGGAGAAATTCTCTGGCATTCAACTCTCTGGCATTC       85039
1022   .................................................  1022
85040  AGAGGTGATAAAGGCTACAGCACGAGGCAATAGGACTGGGACTGTCGTCTCC  85089
1022   .................................................  1022
85090  TTTGGCTGCTGCTGTGCCCCTAAATGTACCCTGCTTTTCCCACCTTTT   85139
1023   .............GTnHisSerGlnThrPro                    1028
85140  CTTTCTCTGTCTGCCCTCACTGTGCCTTGTCCCAGCATTCTCAGACACCA  85189
1029   SerThrLeuAsnProThrMetProAlaSerGluArgThrValAlaTrpVa  1045
85190  TCCACATTGAACCCCACAATGCCAGCCTCTGAGCGGACAGTGGCCTGGGT  85239
1045   lSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHisIleGluA  1062
85240  CTCCAACATGCCTCACCTGTCGCGCTGACATCGAGAGTGCCCACATCGAGC  85289
1062   rgGluGluTyrLysLeuLysGluTyrSerLysSerMetAspGluSerArg  1078
85290  GGGAAGAGTACAAGCTCAAGGAGTACTCCAAAATCGATGGATGGAGCGCCGG  85339
1079   LeuAspArgValLys                                    1083
85340  CTGGATAGGGTACGA                                    85354
```

seq_name: gb_pr3:AF047711

seq_documentation_block:
```
LOCUS       AF047711      4368 bp    mRNA           PRI       14-JAN-1999
DEFINITION  Homo sapiens nGAP mRNA, complete cds.
ACCESSION   AF047711
VERSION     AF047711.1  GI:4105588
KEYWORDS    .
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4368)
AUTHORS     Noto,S., Maeda,T., Hattori,S., Inazawa,J., Imamura,M., Asaka,M. and
            Hatakeyama,M.
TITLE       A novel human RasGAP-like gene that maps within the prostate cancer
            susceptibility locus at chromosome 1q25
JOURNAL     FEBS Lett. 441 (1), 127-131 (1998)
MEDLINE     99093006
REFERENCE   2  (bases 1 to 4368)
AUTHORS     Noto,S., Maeda,T., Hattori,S. and Hatakeyama,M.
TITLE       Direct Submission
JOURNAL     Submitted (11-FEB-1998) Department of Viral Oncology, The Cancer
            Institute, Japanese Foundation for Cancer Research, 1-37-1
            Kami-Ikebukuro, Toshima-ku, Tokyo, Japan
FEATURES             Location/Qualifiers
     source          1..4368
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1q24"
     CDS             126..3545
                     /note="ras GTPase activating protein"
                     /codon_start=1
                     /product="nGAP"
                     /protein_id="AAD04814.1"
                     /db_xref="GI:4105589"
                     /translation="MQTEVPAERSPRRRSISGTSTSEKPNSMDTANTSPF  GFFS
                     KRLKGSIKRTKSQSKLDRNTSFRLPSLRSTDDRSRGLPKLKESRSHESLLSPSCSTVEC
                     LDLGRGEPVSVKPLHSSILGQDFCFEVTYLSGSKGFSCNSASERDKWMENLRRTVQPN
                     KDNCRRAENVLRLWIIEAKDLAPKKKYFCELCLDDTLFARTTSKTKADNIFWGEHFEF
                     FSLPPLHSITVHIYKDVEKKKKDKNNYVGLVNIPTASVTGRQFVEKWYPVSTPTPNK
                     GKTGGPSIRIKSRFQTITILPMEQYKEFAEFVTSNYTMLCSVLEPVISVRNKEELACA
                     LVHILQSTGRAKDFLTDLVMSEVDRCGEHDVLIFRENTIATKSIEEYLKLVGQOYLHD
                     ALGEFIKALYESDENCEVDPSKCSSSELIDHQSNLKMCCELAFCKIINSYCVFPRELK
                     EVFASWKQQCLNRGKQDISERLISASLFLRFLCPAIMSPSLFNLMQEYPDDRISRTLT
                     LIAKVIQNLANFAKFGNKEEYMAFMNDFLEHEWGGMKRFLLEISNPDTISNTPGFDGY
                     IDLGRELSVLHSLLWEVVSQLDKGENSFLQATVAKLGPLPRVLADITKSLTNPTPIQQ
                     QLRRFTEHNSSPNVSGSLSSGLQKIFEDPTDSDLHKLKSPSQDNTDSYFRGKTLLLVQ
                     QASSQSMYSEKDERESSLPNGRSVSLMDLQDTHAAQVEHASVMLDVPIRLTGSQLSI
                     TQVASIKQLRETQSTPQSAPQVRRPLHPALNQPGLQQLPSFQNPVYHLNNPIPAMPKA
                     SIDSSLENLSTASSRSQSNSEDFKLSGPSNSMEDF'KRSTQSEDFSRRHTVPDRHIP
                     LALPRONSTGOAQIRKVDQGGLGARAKAPPSLPHSASLRSTGSMSVVSAALVAEPVQN
                     GSRSRQQSSSSRESPVPKVRAIQROQTQQVOSPVDSATMSPVERTAAWVLNNGQYEED
                     VEETFQNLDEAKHAEKYEQEITKLKERLRVSSRLEEYERRLLVQEQOMQKLLLEYKA
                     RLEDSEERLRRQOEEKDSQMKSIISRLMAVEEELKDHAEMQAVIDAKQKIIDAQEKR
                     IVSLDSANTRLMSALTQVKERYSMQVRNGISPTNPTKLSITENGEFKNSSC"
```

```
BASE COUNT    1292 a    1022 c    1012 g    1042 t
ORIGIN
```

alignment_scores:
```
                Quality:    2584.50        Length:  1303
                Ratio:      2.924          Gaps:    41
Percent Similarity: 67.843   Percent Identity: 45.510
```

alignment_block:
US-09-294-298-6 x AF047711

Align seg 1/1   to: AF047711   from: 1   to: 4368

```
     18  GlyPheLeuSerArgArgLeuLysSerSerIleLeuLysArgThrLysSerGl   34
         |||||:|||:||:.:|||||||||::||||||||||||||||||||||||
    246  GGATTCTTCAGCAGCGCCTGAAGGCTCCATCAAGAGGACCAAAGCCA         295

     34  nProLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheA   51
         |||||||||||:|:||||||||||:::||||||||||||  ||||  :::|
    296  GTCAAAGCTTGACAGAAACACGAGCTTTCGG.......CTTCCATCCCTTC  339

     51  rgSerAlaAspHisAspArgArgAlaArgLeuMetGlnSerPheLysGluSer  67
         |||||:|||  |||||||:||||  ::  ::||||||||||||
    340  GCAGTACAGAT...GACAGGTCTCGTGGGCTGCCTAAACTAAAAGAGTCA   386

     68  HisSerHisGluSerLeuLeuSerProSerSerAlaAlaGluAlaLeuGl   84
         ||||||||||||||||::|||||||||:||||||||||||||||||||
    387  CGTTCCCATGAATCCTTGCTGAGCCCATGCAGCCAGCAGTGAATGTCTGGA  436

     84  uLeuAsnLeuAspGluAspSerIleIleLysProValHisSerSerIleL  101
         ::|||:::|||   |||   ::||||||||||||::||||||||||||||
    437  TCTTGGTAGAGGGAACCTGTATCAGTGAAACCACTTCAGTAGCATCC     486

    101  euGlyGlnGluPheCysPheGluValThrThrSerGlyThrLysCys   117
         |||||||||||||||||||||  |||||||||||:::|||||
    487  TTGGACAAGACTTCTGCTTTGAGGTTACCTACTTAAGTGAAGTAAATGC   536

    118  PheAlaCysArgSerAlaAlaGluArgAspLysTrpIleGluAsnLeuGl  134
         ||||||||||||||||||||||||||||||||||||  |||||:|||||
    537  TTCAGCTGTAATTCTGCTTCGAGAGAGACCAAGTGGATGGAAAACCTTCG  586
```

```
1418 |||||||||||::||||::|||||||||||
     TGAGCTGGCTTTCTGCAAGATCATCAACCTTACTGTGTTTCCCTCGTG 1467

 451 1uLeuLysGluValPheAlaSerTrpAlaArgLeuArgCysAlaGluArgGly 467
     |||||||||:|||||||||||||:|||||||||||||||||||||:|||
1468 AGTTGAAAGAAGTGTTGCATCATGGAAGCAGCAGTGCCTGAACCGTGGC 1517

 468 ArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuArgph 484
     :::::::||||||:::||||||||||||||||||||||||||||||||||
1518 AAGCAAGACATCAGCGAGAGGCTCATCAGTGCCTCATTATTCTCCGTTT 1567

 484 eLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGlnGluT 501
     |||||:|||||||||||||||||||||||||||||||||||||||||||
1568 TCTGTGTCCAGCCATTATGTCTCCCAGTCTTTCAACCTTATGCAGGAGT 1617

 501 yrProAspGluGlnThrSerArgThrLeuThrGluIleAlaLysValIle 517
     |||||:|||||||||||||||||||||||||||||||||||:|||||||
1618 ATCCTGATGACCGCACATCTCGGACTCTAACTCTTATTGCCAAGGTCATT 1667

 518 GlnAsnLeuAlaAsnPheSerLysSerPheThrSerLysGluAspPheLeuGl 534
     ||||||||||||||||||||||||||::||:||||||||||||||||||||
1668 CAGAACCTGGCCAACTTTGCCAAGTTTGGTAACAAGAGGAATACATGC 1717

 534 yPheMetAsnGluPheLeuGluLeuGluGluLeuTrpGlySerMetGlnGlnPheL 551
     ||||||||||||||||||||||||:||||::|||||||||||:||||||||
1718 ATTCATGAATGATTTTTAGAACATGAATGGGGTGGAATGAGCCCTTC 1767

 551 euTyrGluIleSerAsnLeuAspThrLeuHisThrAsnSerSerPheGlu 567
     ||||||||:||||||||||||:::|||:::|||||||::||:|||
1768 TTTTGGAGATCTAATCCAGACACCATCCAAACACCCCAGGCTTTGAT 1817

 567 GlyTyrIleAspLeuGlyArgGlyLeuSerThrLeuHisAlaLeuLeuTr 584
     |||||:|||||||||||||||||||||||||||||::|||||||:::|||
1818 GGTTACATTGATCTGCGGCCCGAGACTTTCAGTTTTGCATTCCTTACTGTG 1867

 584 pGluValLeuProGlnLeuSerLys...............GluA 594
     |||||:||||::||                          |:::
1868 GGAAGTAGTTCCCAACTGATAAGGGTGAAAATTCCTTCCTACAGGCCA 1917

 594 laLeuLeuLysLeuGlyLeuProLeuProArgLeuLeuSerAspIleSerThr 610
     :::::|||||||||||||||||||||||||:|||||||||||||:|||||
1918 CCGTGGCAAAATTGGGCCCTCCCCTGTTCTTGCTGATATTACCAAG 1967

 611 AlaLeuArgAsnAsnProAsn...IleGlnArgGlnProSerArgGlnSerGl 626
     :::::::||||||||:|||   ||:|||||||||||||:|||||:||||
1968 TCATTGACTAATCCTACGCCAATACAAGACAACTGAGACGCTTCACTGA 2017

 626 uArgAlaSerGlnProMetValLeuArgGlyProSerAlaGlyGluMetG 643
     |:::|
2018 ACAT.......................................... 2021

 643 lnGlyTyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerPhe 659
                   ||||||:||
2022 .............AACTCCAGTCCCAAATGTCAGTGGAAGC 2048

 660 MetAlaArgGlyLeuAsn.......................SerSerMetAspMe 670
     :::::::|||||||:::                        :::|||||
2049 CTCTCCCTGGGCTGCCAGAAAATATTTGAAGACCCCACTGACAGTGATTT 2098

 670 tAlaArgLeuProSerProThrLysGluLysProProProProProG 687
     ||:::|||||||:::::||||              ||:|||||||||
2099 GCATAAACTAAAATCCCAGGCAGGACACAC......ACAGACAGTACT 2142

 687 lyGlyGlyLysAspLeuPheTyrValSerArgProLeuAlaArgSer 703
     |:::||||||||||||||||:|:||||||||         |||
2143 TCAGAGGGAAAAACATTATTGCTGGTTCAGCAA.........GCCTCC 2180

 704 SerProAlaTyrCysThrSerSerSerAspIleThrGluProGluGlnIly 720
     |||   ::: ::||||||||||:::||||              |||
2181 TCTCAGACATGACTTATTCTGAAAAGGAT.......GAAAGGGAAAG 2221

 720 sMetLeuSerValAsnLysSerValSerMetLeuAspLeuGlnGlyLysAspG 737
     :                       ||||||||:::::::|||||||
```

```
 134 nArgAlaValLysProAsnLysAspAsnSerArgArgValAspAsnValL 151
     :||:||:||::|||||||||||::||||||||||||||||||||||||
 587 CAGGACAGTTCAACTAATAAGGACAATTGCAGGCGAGCTGAAAATGTTC 636

 151 euLysLeuTrpIleIleGluAlaArgGluLeuProProLysLysArgTyr 167
     |:::||||||||||||||||||::||     ||:||||||:::|||
 637 TTCGTTTATGGATCATTGAAGCCAAGGACCTTGCCCCTAAAAAGAAATAT 686

 168 TyrCysGluLeuCysLeuAspAspMetLeuTyrAlaArgThrThrSerLy 184
     :::||||||||||||||||||||||||||||||||||||||||||
 687 TTCTGTGGAACTGTGCCTTGATGATACCCTTTGCTCGTACAACCAGCCA 736

 184 sProArgSerAlaSerGlyAspThrValPheTrpGlyGluHisPheGluP 201
     |     ||:::   .GACAAAGCA......GACAATATTTTCTGGGGCGAACATTTTGAAT 777
 737 GACCAAAGCA......GACAATATTTTCTGGGGCGAACATTTTGAAT 777

 201 heAsnAsnLeuProAlaValArgAlaLeuArgLeuHisLeuTyrArgAsp 217
     ||  :::||||||   :::||||:::|||||:::|:|:::||
 778 TCTTCAGCCTTCCACCTTCATAGTATCACAGTTCACATTTACAAGGAT 827

 218 SerAspLysLysArgLysLysAspLysAlaGlyTyrValGlyLeuValTh 234
     |||:::||||||:::||||||||||:::||||||||||||||||||
 828 GTGGAAAAAGAAAAAAAAGGACAAGAATAATTATGTAGGGCTAGTCAA 877

 234 rValProAlaAlaThrLeuAlaLeuGlyArgHisPheThrGluGlnTrpTyrP 251
     :::::||||||||||:::||||||:::||||||::::|||||||
 878 CATCCCCACTGCCAGTGCGACTGGTCGCCAATTTGTAGAAAAGTGGTATC 927

 251 roValThrLeuProThrGlySerGlyGlySerGlyGlyLysGlyMetGlySerGly 267
     ||||||
 928 CAGTGACTACACCTACACCCCAC.......................... 950

 268 GlyGlyGlySerGlyGlyGlySerGlyGlyLysGlyGlyGlyGlyGlyCy 284
     |||||||                          ||
 951 ............................AAAGGAAAGACAGGAGG 967

 284 sProAlaValAlaArgLeuLysAlaAlaArgTyrGlnThrMetSerIleLeuProM 301
     |||   ||:::||||:::||||||||||||||:::||||||||
 968 ACCTTCTATTCGGATTAAAATCACGTTCCAAACTAACTCACCATTCGCCTA 1017

 301 etGluLeuTyrLysGluGluPheAlaGlyPheAlaGluThrValThrArgMet 317
     |||||||:::|||||:::|||||||||||||||||||||||||
1018 TGGAGCAATACAAAGAATTTGCAGAATTTGTCACCGACAACTACCACCATG 1067

 318 LeuCysAlaValLeuGluGluProAlaLeuAsnValLysGlyLysGluGluVa 334
     |||||||:::||||||||||:::||||||:::|||||||||||
1068 CTGTGTTTCTGCCTTGACCCAGTAATTAGTGTGAGAAATAAAGAGGAGTT 1117

 334 lAlaSerAlaLeuValHisIleLeuLeuGlnSerThrGlyLysAlaLysAspP 351
     :|||   :::|||||||||:::|||||||||||||||||||
1118 GGCTTGCCTTAGTGCACATTCTTCAAAGTACTGGCAGAGCCAAGGATT 1167

 351 heLeuSerAspMetAlaMetSerGlyValAlaAspArgPheMetGluArgGlu 367
     ||||||:::|||:::|||||||||||||||||||||||||
1168 TTCTGAGTGACTTGGTATGTCTGAGGTGGATCGTGTGGAGCATGAT 1217

 368 HisLeuPheAlaArgGluAsnThrLeuAlaThrLysAlaIleGluGluGluTy 384
     ||||||||||||||||||||||||||||||||||||||||||
1218 GTCTTGATCTTCAGAGAGAACACTATTGCCACCAAATCCATTGAGGAATA 1267

 384 rMetArgLeuIleGluGlyGlnLysTyrLeuLysTyrLeuAspAlaLeuIleGlyGluPheI 401
     :::|||||||||:::||||||||||||||||||||||||||||
1268 CCTCAAGTTGTGTGGACAACAGTATCTTCATGACGCACTGGGGAGTTTA 1317

 401 leArgAlaLeuTyrLysGluGluSerGluGluAsnCysGluValAspProIleLeuLys 417
     ||:::||||||||||||||||||||:::|||||||||||||||
1318 TCAAAGCTTTGTATGATGAGTCCAGAGAACTGTGAAGTGGATCCCAGCAA 1367

 418 CysThrAlaSerSerLeuAlaGluHisGlnHisGlnGlnAlaAsnLeuArgMetCysCy 434
     ||:   :::|||||||||:::||||||||||||||:::||||||||
1368 TGTTCATCTCAGTGAACTGATAGACCATGCAGAGCAACCTGAAAATGTGCTG 1417

 434 sGluLeuAlaLeuCysLysCysLysValValAlaAsnSerHisCysValPheProArgG 451
```

```
2222  TAGCCTTCCTAATGGTCCGGAGCGCTTCCCTCATGGACCTCCAG...GACA  2268
 737  lyProGlyGlyArgLeuAsnSerSerValSerAsnLeuAlaAlaVal       753
      ::::::||||||     :::||||||            :::  |||
2269  CTCATGCTGCTCAAGTGGAGCATGCATCTGTC.....ATGCTGATGTG      2312
 754  GlyAspLeuLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLe     770
      |||  ||||||     ||::|||| |||:::|||       ::::::
2313  CCTATACGCTTGACCGGAAGCCGGAGCGCTTCCATAACCCAGTGGCCAGCAT   2362
 770  uArgProAlaProAlaGlyArgLeuSerGlnGlySerSerIleT          787
      :::                  ::::||           |||     |||
2363  CAAA.........CAGCTGCGGGAAACCCAGAGCACTCCCCAAA           2397
 787  hrAlaAlaGlyMetArg............LeuSerGlnMetGly           797
      ::||| |||                ||:::||||  |||
2398  GTGCACCCCAAGTGAGAAGGCCCCTGCACCCACCGCCTGAACCAGCCAGT     2447
 798  ValThrThrAspGlyValProAlaGlnGlnLeuArgIleProLeuSerPh     814
      ||:::|             ||||||||
2448  .........GGCCTTCAG............................CCCTTGTCGTT  2467
 814  eGlnAsnProLeuPheHisMetAlaAlaAlaAspGlyProGlyProProAlaG    831
      ||||||||::::||:::|||::  ||       ||| :
2468  CCAGAACCCTGTCTATCACCCTCAATACCCAATTCCAGCAATGCCAAAGG    2517
 831  lyHisGlyGlySerGlyHisGlyHisGlyProProSerHisHisHisHis    847
      |||||||||||
2518  CCTCTATAGAATTCCAGT.........................           2534
 848  HisHisHisHisHisArgGlyGlyGlyGluProProGlyAspThrPheAL    864
      ||:::||||||||
2535  ......................TTGGAGAACCTAAGCACT.......        2552
 864  aProPheHisGlyTyrSerLysSerGluAspLeuSerThrGlyValProL    881
      |||::||||||||
2553  .........GCCAGTTCC......AGAGCCAAAGTAACAGTGAAGACTTC    2588
 881  ysProProAlaAlaSerIleLeuHisSerHisSerTyrSerAspGluPhe    897
      :::||| ::::||||||||
2589  AAGCTCAGTGGACCCAGCAATAGCAGCATGGAAGATTTCACTAAACGTAG    2638
 898  .........GlyProSerGlyThr.......AspPheThrArgArgGl      908
      ||||||||::::||:::
2639  CACTCCAGAGTGAGGACTTCTCCAGGCGGCCACACGGTGCCAGATAGACACA   2688
 908  nLeuSerLeuGlnAsp.......AsnLeuGlnHisMetLeuSerProProGlnI   924
      :::||||  ::::/||   :::||||
2689  TACCTCCTTGCTTTGCCACGACAAAAATAGTACTGGGCACGGCCCAGATCCGA   2738
 924  leThrIleGlyProGlnArgProAlaProSerGlyProGlyGlyGlySer    940
      ||  :::||||     |||||
2739  AAAGTGGACCAGGTGGGTTAGGTGCCCGAGCCAAAGCCCCCACCATCCCT    2788
 940  GlyGlyGlySerGlyGlyGly.....GlyGlyGlnProProProLe       955
      |||||||||||     ::|| |||::::|
2789  GCCACACAGTGCTTCTTTACGTAGCACCGGACGCATGTCAGTGGTGTCCG    2838
 955  uGlnArgGlyLysSerGlnLeuThr................ValSerA      967
      ::::||||| ||| ||:::|||          |||||
2839  CAGCC.................................CTGGTGGCCGAA   2855
 967  laAlaGlnLysProArgProSerGlyAsnLeuGlnLeuGlnSerProGlu    983
      |||||
2856  CCTGTGCAGAATGGGAGCCGGTTCCCGGCAGCAGTCCTCTTCCTCCAGAGA   2905
 984  ProSerTyrGlyProAlaArgProArgGlnGln...SerLeuSerLysGl   999
      ::: ::||| |||||||::| ||| :::::||:::::||
2906  GAGCCCTGTT.....................................      2915
 999  uGlySerIleGlyGlySerGlyGlySerGlyGlyGlyGlyGlyGlyGlyL   1016
```

```
1016  euLysProSerIleThrLys...GlnHisSerGlnThrProSerThrLeu   1031
      :::|||       ::|||  |||||
2916  ....CCCAAAGTTAGAGCAATCCAGAGACAACAGACACCAGAGGTTCAG    2960
1032  AsnPro.........ThrMetProAlaSerGluArgThrValAlaTr     1044
      ::::|||       |||||             :::||||  ::::||||
2961  TCACCTGTGGACTCTGCCACAATGTCCCCAGTAGAGAGGACAGCAGCCTG   3010
1044  pValSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHisIleG   1061
      :::                  |||::::||||
3011  GGTTCTGAACAATGGCAGTATGAAGAGGATGTGGAA.......         3047
1061  luArgGluGluTyrLysLeuLysGluTyrSerLysSerMetAspGluSer   1077
      :::  |||          |||    :::||||||
3048  .....................GAAACTGAGCAAAATCTAGATGAAGCC    3074
1078  ArgLeuAspArgValLysGluTyrGluGluLeuHisSerLeuLysGl      1094
      :::                  :::||||||
3075  AAG.......CATGCTGAGAAGTATGAACAAGAAATTACTAAACTGAAGGA   3118
1094  uArgLeuHisMetSerAsnArgLysLeuGluGluTyrGluArgArgLeuL   1111
      ||:::||||||::::||:::||   |||||||||||||||||||||||||
3119  GCGCCTGAGAGTTTCCAGCCGGCCACTCGAGAATATGAACGCCCGCTTGC   3168
1111  euSerGlnGluGluGlnThrSerLysIleLeuMetGlnTyrGlnAlaArg   1127
      ||  ||       ::::::|||    ::||  :::||||
3169  TGGTGCAGGAGCAGCAGATGCAGAAGCTGCTGCTGGAATACAAGGCCCGA   3218
1128  LeuGluGlnSerGlyLysArgLeuArgGlnGlnGlnValGluLysAspSe   1144
      |||||||||  |||           |||||
3219  CTGGAGGACAGCGAGGAGCGGCGTCCGAAGACAGCAGGAAGAAAAGATAG    3268
1144  rGlnIleLysSerIleIleGlyArgLeuMetLeuValGluGluGluLeuA   1161
      |||::||||||||||||||     ::||||||  |||||||||||
3269  CCAGATGAAAAGCATCATCAGCAGGCTAATGGCTGTGGAAGAGGAACTGA    3318
1161  rgArgAspHisProAlaMetAlaGluProLeuProGluProLysLysArg   1177
      ::::||||     ||             |||||::||||||
3319  AGAAGGATCATGCTGAGATG..CAAGCAGTTATTGATGCAAAGCAGAAA    3365
1178  LeuLeuAspAlaGln.ArgGlySerPheProProTrpValGlnGlnThrA   1194
      :::||||||||:::|||||||                 ||||::::|||
3366  ATAATTGATGCACAGAAAAACGGATCGTGTCCCTGGATTCAGCCAACAC   3415
1194  rg........ValArgTrpPro......HisLeuGlyThrAlaTrp...   1204
      :::|||       :::|||||||||
3416  CAGACTGATGAGGCGCGTGACCCAAGTGAAGGAGGCTGTACAGCATGCAGG   3465
1205  ....................ProProGlnProHisProProHisProGlyCysAr   1216
      :::                  |||::||       ||
3466  TCCGCAATGGCATCTCCCCACCAACCCCACCAAGCTTTCCATCACCGGAG   3515
1216  g.........SerGlnAsnArgThrAlaSerSerGlyThr............  1225
      |||                  :::||
3516  AATGGTGAATTCAAAAACACAGCACGTGCTGACGGGCTTTGTCTGTGGAAGG   3565
1226  ....................ProGlnThrThrSerProProSerIleThrAsp   1236
      :::                  ::::||||||
3566  AGACAGAAGGAAATTGACCCACTCTCCTATCTCCA...........GAC    3603
1237  LeuLeuProCysAlaProTyrProGlyProProSerValThrAspLeuLe   1253
      |||                  :::|||||
3604  CTT...............TACCTCTAGCCCCTCCAGGTTTACAGAATGTTGC   3638
```

```
1253  uProSer  1255
      |||
3639  TACTTCA  3645
```

```
DEFINITION   Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
             pieces.
ACCESSION    AC019800
VERSION      AC019800.1  GI:6665097
KEYWORDS     HTG; HTGS_PHASE2.
SOURCE       fruit fly.
ORGANISM     Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
             Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
             Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1  (bases 1 to 56366)
AUTHORS      Adams,M. and Venter,J.C.
TITLE        Direct Submission
JOURNAL      Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
             Rockville, MD, USA
COMMENT      This sequence was identified as CDM.10210991 by the submitter.
             For more information on this record e-mail to fly@celera.com.
             * NOTE: This is a 'working draft' sequence.
             * This sequence will be replaced
             * by the finished sequence as soon as it is available and
             * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1..56366
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
BASE COUNT     16234 a   11924 c   11808 g   16400 t
ORIGIN

alignment_scores:
          Quality: 1257.50           Length:  1432
          Ratio:   1.642             Gaps:      57
Percent Similarity: 53.492    Percent Identity: 29.190

alignment_block:
US-09-294-298-6 x AC019800/rev

Align seg 1/1   to reverse of: AC019800  from:  1   to: 56366
```

```
25776  TTTGATCTACTCCGTCGCCAAGTCGATCGCTCCGAATGCAGAGCACACG  25727
   19  FheLeuSerArgArgLeuLysSerSerIleLysArgThrLysSerGlnPr   35
  145  ArgArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGluLe  161
26122  TTCCTTGCCACAGA...TCCAATCCTTGAAGCGGACCAAGTCGGTGAC  26076
25726  CGTCGCACGGACAACTCGCTGAAGATGTGGTGACGAGCGGAAAAATCT  25677

  161  uProProLysLysArgTyrTyrCysGluLeuCysLeuAspAspMetLeuT  178
   35  OLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheArgS   52
26075  CAAGCTGGAGCGGACCAAG...                             26057
25676  GCCGCCCAAGAAGCGTTACTTTTGCGAACTGCAATTGGACAAGAGCGCTGT  25627

  178  yrAlaArgThrSerLysProArgSerAlaSerGlyAspThrValPhe    194
   52  erAlaAspHisAspArgAlaArgLeuMetGlnSerPheLysGluSerHis   68
26056  ...CGCGGATCCGGCGACTGAGGGGGGCTCCCGC  26027
25626  ACGGCCGGACTTCGGTGAAGCTGCAGACG.......GATCTGCTGTTT   25586

  195  TrpGlyGlnHisPheGluPheAsnAsnLeuProAlaValArgAlaLeuAr  211
   69  SerHisGluSerLeuSerProSerSerAlaAlaGluAla........    82
26026  TCGCACGAGAGTCTGCTGTCCAGTCACGCCGTCCATGTCCACCATAGGTGA  25977
25585  TGGGGGAGCGACTTCGATTCCCCGACATACCCGAGATTAATGTGATCAC  25536

  211  gLeuHisLeuTyrArgAspLysSerLysArgLysLysAspLysLysAlaG  228
   82  .....................................             82
25535  TGTTAACGTTTTCCGTGAGGTGGACAAGAAGAAGGCGGACAAATACC  25486

  228  lyTyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgHis  244
   82  .....GTTCTCGGCTCATAATGGAGTTTTGGCCAAACCCTCAAACCATTTCTTTT  25927
25485  AATTTGTGGGATCCGTGAAGATACCCGTGCACGATGTCACCTTCCAGATTG  25436

  245  PheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGlySe  261
   83  ...........LeuGluLeuAsnLeuAspSerIleIleLeuLysPro    95
25877  CCAAGAGCATGGAGGCAATTCCTCAAGCTGACGGGCGAACAGTATCTGCAG
25435  CCCTGCGAACAATGGTATCCCATA...CTGAGCGACAAGGCGGGCGACAG  25389

  261  rGlyGlyMetGlyGlyGlyLyGlyGlyGlyLySerGlyGlyLySerGlyG  278
   96  ValHisSerSerIleLeuGlyGlnPheCysPheGluValThrThrSe    112
25826  GTGCATCAGTCGGTTCGGGACGCGTCGTCACTGTTTCCAGGTACGGGGGCGG  25876
25388  TCTGGGCAGGACATCGGGC...                             25354

  278  lyLysGlyLysGlyGlyCysProAlaValArgLeuLysAlaArgTyrGln  294
  112  r...SerGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAspL  128
24956  CGAGTGGATCCACCAAGAGCATAAACGAGCTAATTCAGTCGGAGAGGACTG
25353  CCAAGGACAAGGAGCAATTGCCCACGCTGCAGGATCAAGTGTCGTTTCCAG  25304

  295  ThrMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyrVa  311
  128  ys........CysThrAlaSerSerLeuAlaG   144
24906  GACACGAGCAGGCCGCCGCTTGCGTCCGCGGCGGGGCGTGGCAGTGCATC
25303  AGCACGACATCCTGCCCATCAATGTGTACGGCAACTTTTTGACGTACCT    25254

  311  lThrAsnHisTyrArgMetLeuCysAlaValLeuGlnProAlaLeuAsnV  328
25253  CAAGGAGAACTATAAGCGCGTGTGCGAGACCCTGGAGCCGGGTCATCGGAG  25204

  328  aLLysGlyLysGluGluValAlaSerAlaAlaValHisIleLeuGlnSer  344
25203  TCAAGGCCAAGGAGGACATTGGACAGGCACTGGTGCTGCTGATGCACGCA   25154

  345  ThrGlyLysAlaLysAspPheLeuSerArgSerMetAlaMetSerGluVaLAs  361
25153  CAGGGATTGGCGGGCGCCCTTCCTCACCGATGTGGTGGCCCTCGATCTGCT   25104

  361  pArgPheMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaT  378
25103  GCGAGTTGGCGAT...CAGAAGCTTACGTTCAGGGGCAACTCCTTGGCCA   25057

  378  hrLysAlaIleGluGluTyrMetArgLeuIleGlyLysTyrLeuLys    394
25056  CCAAGAGCATGGAGGCATTCCTCAAGCTGACGGGCGGACAGTATCTGCAG   25007

  395  AspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsnCy  411
25006  GACACACTATCGCGACCAATAAACGAGCTAATTCAGTCGGAGAGGACTG   24957

  411  sGluValAspProIleLys........CysThrAlaSerSerLeuAlaG  425
24956  CGAGTGGATCCACCAAGAGCATAAACGAGCTTCGTCCGCGGCGGGCGACC   24907

  425  luHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysVal  441
24906  GACACGAGCAGGCCGCCGCTTGCGTCCGCGGCGTCGGAGGGGCGTGGCATC   24857
```

```
  693  eTyrValSerArgProProLeuAlaArgSerSerProAlaTyrCysThrS  710
                                                |||||||||:::|
24051                                          GCTTACTGCAGCA  24039

  710  erSerSerAspIleThrGluProGlnLysMetLeuSerValAsnLys  726
       |||||||||| ||||:::::               ::         |
24038  GCAGCAGCAGCAGCAGCAGC.                 AGCAGCAACAAC  24007

  727  SerValSerMetLeuAspLeuGlnGlyAspGlyProGlyGlyArgLeuAs  743
       ||| |||  :::                                   ||
24006  AGCAACAGCAACATC.                             AGCAGCTGCAA  23981

  743  nSerSerValSerAsn......LeuAlaAlaValGly..AspLeuLe  757
       |||||| :::::           ::::::           :::::||
23980  CAGCATGGCCATCAGCAACAGCCGCCACCACCAGCATCCACTCCAGATGCT  23931

  757  uHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaP  774
       ::||||:||::::||||               |||::
23930  CTCCAGCATTCACAAACCTCCATT...GCCGGCAACCAATATGAGTTCGC  23884

  774  roAlaGlyArgLeuSerGlnGlySerSerIleThrAlaAlaGly  790
       :|:||||::||:|||  ::|||::||||||||||||||||||||
23883  CAGGAGGCCTGCAGCATGCCCAATCCAGACCTCGATGGCGTCCTCATCG  23834

  791  MetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnGl  807
       :::
23833  CTT.                                               23831

  807  nLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaA  824
23831  .                                                  23831

  824  spGlyProGlyProAlaGlyHisGlyGlySerGly.....His  838
       ||:                                    |||
23830  AATGGAGCAGCAGCAATTTGCTGCAC  23804

  839  GlyProProSerHisHisHisHisHis............Hi  851
       |||          |||||||||| ||||||
23803  GGC.        CACCAGCAGCATGCCCATCACCCGCAGCAACTGCA  23766

  851  sHisHisArgGlyGlyGluProProGlyAspThrPheAlaProPheHisG  868
       |    |||:|                 :
23765  TCCCACATCACTGC.            CCGCCG.                  23747

  868  lyTyrSerLysSerGluAspLeuSerThrGlyValProLysProProAla  884
                                            ::|||:|  |||
23746                                      GCGCCACAGACCAGTGCC  23729

  885  AlaSerIleLeuHis............SerHisSerTyrSerAspGl  896
       ::|||                      ::: |||
23728  TCCAGCACTATGGAGCCCATGGATCCGCATGACTATCCGTATATGTCGCA  23679

  896  uPheGlyProSerGlyThrAspPheThrArgArgGlnLeuSerLeuGlnA  913
       :               :|||||:
23678  C.              AATGGCAATGACTACGAG.                 23660

  913  spAsnLeuGlnHisMetLeuSerProProGlnIleThrIleGlyProGln  929
                                 |||  |||
23659                           ACCAGCACGCCTTCGAGCACCTCCAGGACA  23627

  930  ArgProAlaProSerGlyPro.....GlyGlyGlySerGlyGlyGlySe  944
       |||                       |||
23626  CTGCCACGGAATGGAAATCCCAATGCCAATGGCAACGTGGGCAGCAGCAA  23577

  944  rGlyGlyGlyGlyGlyGlnProPro.LeuGlnArgGlyLysSer  960
       ::||||||        :::|||||
23576  CAATAACCAGAGCGGCAGCTACGATGACATGCACGGGGAGTTCCAAATCC  23527

  961  GlnGlnLeu........ThrValSerAlaAlaGlnLysProAr  972
       :::|||
23526  AGATCTCTGGGTTCGATACGAGCAGTGCTTTTGTCTGCAAGTCGCCCACA  23477

  972  gPro.......Ser.SerGlyAsnLeuLeuGlnSerPro  982
```

```
  442  ValAsnSerHisCysValPheProArgGluLeuLysGluValPheAlaSe  458
       |||||||| ||||| |||||: :::|| ||||||
24856  TTCGAATCGCACAAGCATTTCCCGCCCCAGTTACGAAATTGCTTTGCGAC  24807

  458  rTrpArgLeuArgCysAlaGluArgGlyLysArgGluAlaAspArgL  475
       :::|||  |||  :: |||||||||||| ||||||||||||:||
24806  GTTCCGGGAGCGGCTTGCAGCAGCGGCCGTCAGGATATGGCCGACAACC  24757

  475  euIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSer  491
       ||||||||||||||||||||| |||:|| :::
24756  TGATCTCGGCGAGCATTTTCCTCGCGCTTCCTCGTGCCCGGCCATCCTGTCG  24707

  492  ProSerLeuPheGlyLeuMetGlnGluTyrProAspGlGlnThrSerAr  508
       ||||||||||| |||:|| :::|| ||||||::::::||
24706  CCGTCGCCTTTCAATATCACCAGCGGACTGCGTCCGCACGGGCCACCCG  24657

  508  gThrLeuThrIleAlaLysValIleGlnAsnLeuAlaAsnPheSerL  525
       |:::||||||||::|||||:::|||||| ||||||||||||||::::
24656  CAATCTCACACTGCTGGCCAAGACCCTGCAAGACCATTGGCCAATTTCACCC  24607

  525  ysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGlu  541
       :||||:||::||||:::|| ||||||||||||||||
24606  GCTTCCAGGGCAAAGAGAACTTTATGAAGTTCTCAACGATTTCCTTCGAG  24557

  542  LeuGluTrpGlySerMetGlnGlnPheLeuGluTyrGluIleSerAsnLeu..  557
       ||| :::  |||||||||||||||||||||::::
24556  CAGGAGGCCGCCTCGCCTCGCAACAGTTTCTGGAGGATTATATCCACGGCGCC  24507

  558  .......AspThrLeuThrAsnSerSerSerPheGluGluGlyT  569
       ||:::::     :|||  |||||||
24506  GGAGCACCCAGCTCCAGACTCGATCCTCGAT.........TGGGCCGGGT  24466

  569  yrIleAspLeuArgGluLeuSerThrLeuHisAlaAlaLeuTrpGlu  585
       ||||||:: :::||||||||  ||||||||||||:|||
24465  ACATCGACCAGGGCAAACAGTTGTCCATACTACACAGTTTGCTCAGCGAA  24416

  586  ValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGlyProLeuPr  602
       :::::||                  ||:|||
24415  AGCCTGGCCAAGCTCGCCGAGGCCAGGCCAGCACGAGCTGGATCCGTTGCA  24366

  602  oArgLeuLeuSerAspIleSer............ThrA  611
       ::::|||:::::::||||||||
24365  GCATATTCTCGATGAAATCAGCCGAGCCAAAGAGCATGGCATGGCAGAG  24316

  611  laLeu.........................  612
       ||||
24315  CACTGCGCCGGGTGGATATTTGCCGCCACCTGCTACGCCACTCGATAGCC  24266

  613  .....ArgAsnPro.........................  615
            |||||||
24265  AGCGAGAAATCAGGAGAATCCGCAATCCGGGATCTCCGGGCTCGCACGCTGG  24216

  616  ...AsnIleGlnArgGlnInProSerArgGlnSerGluArgAlaArgSerG  631
       |||   ::|||||:| :::::
24215  CTCCAACTCCGGAGCAGTTACTGCCAACAACAAAGCCAGTTGGCCCAGCCGC  24166

  631  lnProMetValLeuArgGlyPro...SerAlaGlu........MetGln  643
       ::::::||||||||||||||||        :::::   |||::
24165  AGCATGCGGATTGTTAGTAAACCATTATCTGCGGAGCCGCGGCCATCATGCGA  24116

  644  Gly.TyrMetMetArgAspLeuAsnSerSerIleAspProGlnSerPheM  660
       |||:||::::|||   |||  |||||:::::::::::::
24115  GGAGTACTTACGGCCAATTCTCTGGAGAGAATATCTTTAGATACAATGA  24066

  660  etAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerPro  676
       :::::||| :::
24065  CTCCAACTCCGAGCAGTTACTGCCAACAACAAAGCCAGTTGGCCCAGCCGC  24052

  677  ThrLysGluLysProProProProProGlyGlyLysAspLeuPheAspLeuProH  693
       ::::|||  |||  ||||||||||||::::::
24052                                                       24052
```

```
           :||                                                   ::  ||||||::
23476 CCCATGATGAAATCCAGTTTGGACCAGCGGGGGCCGGGGACGAAGTCATCA 23427

 983 GluProSerTyrGlyProAlaArgProArg............. 992
           :::  :::||| |||
23426 CAAACTGAATTTGGGAATACCCGATCACTCAGGTGGCTATGTGCGGGGTA 23377

 993 ......GlnGlnSerLeuSerGluGlyGlySerIleGlyGlySerGlyG 1007
           ::::::||||||||||||
23376 ATAATTTGAATCCAACTGGAATATGCCCAAGAACTTGGAGGATCTGGAC 23327

1007 lySer............GlyGlyGlyGlyGlyGlyLeuLysPro 1018
           |||        ||||| ::|||
23326 GATCTGTTCAAGTGACTACGCCGGAGGAGCATGACGTGGCCGAACCAGCGAACCA 23277

1019 SerIleThrLysGlnHisSer..GlnThrProSerThrLeuAsnPro. 1033
           ||| |::::||||::||| ||   ||:|||
23276 TCACAATCACACCAGGGTCAGCAGAACCACCCAGGGACATCTGAAGCCGG 23227

1034 .ThrMetProAlaSerGluArgThrValAlaThrTrpValSerAsnMetPro 1049
           ::::::|||:::|||:::
23226 CCGCGGTTCCCGGCAAGGAGCAGCGTCTCGGCGAAAAGCAGT...... 23186

1050 HisLeuSerAlaAspIleGluSerAlaHisIleGluArgGluGluTyrTyrLy 1066
           |||:           ::||||
23185 CACTGCAGTTCTGGCTACCAGAGCCATCTCCACAAATCCC.......... 23147

1066 sLeuLysLysGluTyrSerLysSerMetAspGluSerArgLeuAspArgValL 1083
           |||      ::||::|||
23146 ......TCGCCCTGCAGTCCTCCAGTCCCGTGGAG........ 23117

1083 ysGluTyrGluGluLeuLysLysLeuLysSerLeuLysGluArgGluHisMetSer 1099
           ||
23116 ............AGCCAGCTGAAGGCCCCGGATGGGCAGTCAC 23087

1100 AsnArgLysLeuGluGluTyrGluArgArgLeuuSerGlnGluGluGl 1116
           |||  ||    ::  :: |||  |
23086 AATGCGCCCGCTG...GCCTTCAAGAATCCCTCCTATCAGCTTCAGCCCCA 23040

1116 nThrSerLysIleLeuMetGlnThrTyrGlnAlaArgLeuGluGlnSerGluL 1133
           ||||:::||| ||:::||
23039 AACTGGCTCGTCCAGATCATCCGCACAGAGTAATACACACCAGCAGCAGC 22990

1133 ysArgLeuArgGlnGlnGlnValGluLysAspSerGlnIleLysSerIle 1149
           :|||||||||          |||
22989 AGCAGCAGCAACAACAGTT......GCCAGCCGCTTGAAACCAATT 22946

1150 IleGlyArgLeuMetLeuValGluValGluGluLeuArgArgAspHisProAl 1166
           |||  ||:::        :: ||
22945 GGCGGCCGACTGGTGGCCGGGCGGCTTTCCTCAACAGTGGCGGAGC 22896

1166 aMet.AlaGluProLeu............ProGluProLys......... 1175
           |:: ::||||||     |||  |||::
22895 CTTGGAGGCGCCACTTGACGCCCAGCTCCTCGGACGACAGCTCTCGG 22846

1176 ...........LysArgLeuLeuAspAlaGlnAlaGlnArgGlySer 1185
           ::||||||       ::|||
22845 CGGATAATTACTTCAGTTATGCAGCGGCTGCAGCGTGCTGAGCGGGTATT 22796

1185 rPhePheProProValGlnGlnThrArgValArgTrpProHisLeuGlyT 1202
           ||||||||::    ||  :: ||:::|||
22795 GCCACCAAATTGGAGGCTCAACGCTCGC..........TCAGCGCCGGCAG 22755

1202 hrAlaTrpProProGln............. 1207
           :|:| |||||||||
22754 TAGCTCCTCCACCTCAGCATCTGCGTCCACCTCGAATCTTGGCAAGAGAGCG 22705

1208 ..ProHisProProHisProGlyCysArgSerGlnArgThrAlaSerSer 1223
           ||||||            ::||
22704 GCGGCTCATCCGCCCTACGGGCGGCTCGAATGGCCCGCTGAAGCGCGAGGAT 22655

1223 rGlyThrProGlnThrThrSerProPro...........SerIleThrA 1236
           ::|||           :::::::::
```

```
22654 GTCTACGGCAGTGGCTACGGCGGCAGCAGCAGCAGTGGAAATGTGGCTATGGCTT 22605
           ::     ||||||||::
1236 spLeuLeuPro.Cys......AlaProTyrProGlyProProSerValTh 1250
           ||||| |||           ::|||
22604 GTCCACTTCCAGTGCCGGGGACACCATCAACATCCCACCAGCAGCAGC 22555
           :::           ||||||::
1250 rAspLeuProSerAlaProAspProGlyThrSerProThrThrArgT 1267
           :::|:          |||||||||
22554 AGAATCCGATCGCAGCAGCAGCAGCAAAGGGAACGGGATCAGGAACACAAG 22505
           ::::::          |||||||||||
1267 hr...........GlyArgHisGlnGlyThrAlaGlyLeuSerPro 1278
           :|              |||   :::|||
22504 CAGTATGCCGGCAGTGTGGCCGGGCAGCGTGGGATCGGCCACATCAGCCGC 22455
           ||||||           ::::|||
1279 LeuThrProPro........TrpGlyThrHis 1286
           |||            ||::|||
22454 TCAGAGGCGCCTGAGCTTGGACTCCGCGCCGCCAC 22422
```

seq_name: gb_htg4:AC012162

us-09-294-298-6.rge

Fri Jan 19 06:20:05 2001

```
*    537  1004:  contig of 468 bp in length
*   1005  1084:  gap of unknown length
*   1085  1958:  contig of 874 bp in length
*   1959  2038:  gap of unknown length
*   2039  2248:  contig of 210 bp in length
*   2249  2328:  gap of unknown length
*   2329  2858:  contig of 530 bp in length
*   2859  2938:  gap of unknown length
*   2939  3325:  contig of 387 bp in length
*   3326  3405:  gap of unknown length
*   3406  4165:  contig of 760 bp in length
*   4166  4245:  gap of unknown length
*   4246  4716:  contig of 471 bp in length
*   4717  4796:  gap of unknown length
*   4797  5373:  contig of 577 bp in length
*   5374  5453:  gap of unknown length
*   5454  6050:  contig of 597 bp in length
*   6051  6130:  gap of unknown length
*   6131  6720:  contig of 590 bp in length
*   6721  6800:  gap of unknown length
*   6801  7461:  contig of 661 bp in length
*   7462  7541:  gap of unknown length
*   7542  8104:  contig of 563 bp in length
*   8105  8184:  gap of unknown length
*   8185  8505:  contig of 321 bp in length
*   8506  8585:  gap of unknown length
*   8586  9227:  contig of 642 bp in length
*   9228  9307:  gap of unknown length
*   9308 10091:  contig of 784 bp in length
*  10092 10171:  gap of unknown length
*  10172 11099:  contig of 928 bp in length
*  11100 11179:  gap of unknown length
*  11180 11747:  contig of 568 bp in length
*  11748 11827:  gap of unknown length
*  11828 12767:  contig of 940 bp in length
*  12768 12847:  gap of unknown length
*  12848 13398:  contig of 551 bp in length
*  13399 13478:  gap of unknown length
*  13479 14007:  contig of 529 bp in length
*  14008 14087:  gap of unknown length
*  14088 14867:  contig of 780 bp in length
*  14868 14947:  gap of unknown length
*  14948 15562:  contig of 615 bp in length
*  15563 15642:  gap of unknown length
*  15643 16859:  contig of 1217 bp in length
*  16860 16939:  gap of unknown length
*  16940 17986:  contig of 1047 bp in length
*  17987 18066:  gap of unknown length
*  18067 19341:  contig of 1275 bp in length
*  19342 19421:  gap of unknown length
*  19422 20926:  contig of 1505 bp in length
*  20927 21006:  gap of unknown length
*  21007 22153:  contig of 1147 bp in length
*  22154 22233:  gap of unknown length
*  22234 23236:  contig of 1003 bp in length
*  23237 23316:  gap of unknown length
*  23317 24607:  contig of 1291 bp in length
*  24608 24687:  gap of unknown length
*  24688 25620:  contig of 933 bp in length
*  25621 25700:  gap of unknown length
*  25701 26300:  contig of 600 bp in length
*  26301 26380:  gap of unknown length
*  26381 27488:  contig of 1108 bp in length
*  27489 27568:  gap of unknown length
*  27569 28215:  contig of 647 bp in length
*  28216 28295:  gap of unknown length
*  28296 29440:  contig of 1145 bp in length
*  29441 29520:  gap of unknown length
*  29521 30587:  contig of 1067 bp in length
*  30588 30667:  gap of unknown length
*  30668 32042:  contig of 1375 bp in length
*  32043 32122:  gap of unknown length
*  32123 33740:  contig of 1618 bp in length

*  33741  33820:  gap of unknown length
*  33821  35375:  contig of 1555 bp in length
*  35376  35455:  gap of unknown length
*  35456  36258:  contig of 803 bp in length
*  36259  36338:  gap of unknown length
*  36339  38003:  contig of 1665 bp in length
*  38004  38083:  gap of unknown length
*  38084  39509:  contig of 1426 bp in length
*  39510  39589:  gap of unknown length
*  39590  40719:  contig of 1130 bp in length
*  40720  40799:  gap of unknown length
*  40800  42453:  contig of 1654 bp in length
*  42454  42533:  gap of unknown length
*  42534  43841:  contig of 1308 bp in length
*  43842  43921:  gap of unknown length
*  43922  45410:  contig of 1489 bp in length
*  45411  45490:  gap of unknown length
*  45491  46782:  contig of 1292 bp in length
*  46783  46862:  gap of unknown length
*  46863  48570:  contig of 1708 bp in length
*  48571  48650:  gap of unknown length
*  48651  51042:  contig of 2392 bp in length
*  51043  51122:  gap of unknown length
*  51123  52760:  contig of 1638 bp in length
*  52761  52840:  gap of unknown length
*  52841  54712:  contig of 1872 bp in length
*  54713  54792:  gap of unknown length
*  54793  56521:  contig of 1729 bp in length
*  56522  56601:  gap of unknown length
*  56602  58938:  contig of 2337 bp in length
*  58939  59018:  gap of unknown length
*  59019  60837:  contig of 1819 bp in length
*  60838  60917:  gap of unknown length
*  60918  62988:  contig of 2071 bp in length
*  62989  63068:  gap of unknown length
*  63069  65583:  contig of 2515 bp in length
*  65584  65663:  gap of unknown length
*  65664  68774:  contig of 3111 bp in length
*  68775  68854:  gap of unknown length
*  68855  71235:  contig of 2381 bp in length
*  71236  71315:  gap of unknown length
*  71316  74113:  contig of 2798 bp in length
*  74114  74193:  gap of unknown length
*  74194  77405:  contig of 3212 bp in length
*  77406  77485:  gap of unknown length
*  77486  80936:  contig of 3451 bp in length
*  80937  81016:  gap of unknown length
*  81017  83357:  contig of 2341 bp in length
*  83358  83437:  gap of unknown length
*  83438  87456:  contig of 4019 bp in length
*  87457  87536:  gap of unknown length
*  87537  91532:  contig of 3996 bp in length
*  91533  91612:  gap of unknown length
*  91613  97143:  contig of 5531 bp in length
*  97144  97223:  gap of unknown length
*  97224 105593:  contig of 8370 bp in length
* 105594 105673:  gap of unknown length
* 105674 112071:  contig of 6398 bp in length
* 112072 112151:  gap of unknown length
* 112152 121011:  contig of 8860 bp in length
* 121012 121091:  gap of unknown length
* 121092 143113:  contig of 22022 bp in length
* 143114 143193:  gap of unknown length
* 143194 143587:  contig of 394 bp in length
* 143588 143667:  gap of unknown length
* 143668 143927:  contig of 260 bp in length
* 143928 144007:  gap of unknown length
* 144008 144842:  contig of 835 bp in length
* 144843 144922:  gap of unknown length
* 144923 145360:  contig of 438 bp in length
* 145361 145440:  gap of unknown length
* 145441 145915:  contig of 475 bp in length
* 145916 145995:  gap of unknown length
```

```
*   145996   146532: contig of 537 bp in length
*   146533   146612: gap of unknown length

alignment_scores:
  Quality:   1257.50        Length:      1432
  Ratio:     1.642          Gaps:          57
  Percent Similarity: 53.492   Percent Identity: 29.190

alignment_block:
US-09-294-298-6 x AC012162

Align seg 1/1   to: AC012162   from: 1   to: 191504
```

```
245 PheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGlySe 261
    |||||||||||||||::||     |||
113730 CCCTGCGACAACATGGTATCCCATA...CTGAGCGACAAGCGGCGACAG 113776

261 rGlyGlyMetGlySerGlyGlyGlyGlyGlyGlyGlySerGlySerGlyG 278
    ||||||                         |||||||||
113777 TCTGGCAGGACATCGGGC..............GGCGGCGGCAGTGGGT 113811

278 lyLysGlyLysGlyGlyGlyCysProAlaValArgLeuLysAlaArgTyrGln 294
    :|||  |||   |||::   :::|||:::|||:::|||
113812 CCAAGGACAAGGAGGACAATTGCCCACGCTGAGGACCAAGTGTCGTTCCAG 113861

295 ThrMetSerIleLeuProMetGluGluTyrLysGluAlaGluThrTyrVa 311
    :::       :::|||  :::|||    :::|||
113862 AGCACCGACATCCTGCCCATCAATGTACGGCAACTTTTGACGTACCT 113911

311 lThrAsnHisTyrArgMetLeuCysValAlaValLeuGluProAlaLeuAsnV 328
    :  :::||||||::|||   :::||||||||||||:::
113912 CAAGGAGAACTATAAGCCGCGTGCGAGACCCTGCCGGTGATCGGAG 113961

328 alLysGlyLysGluLeuAlaSerAlaLeuValHisIleLeuGlnSer 344
    :|||  :::|||:::|||:::||:::   :::|||
113962 TCAAGGCCAAGGAGGACATTGGACAGGCACTGGTGCTGCTGATGCACCCA 114011

345 ThrGlyLysAlaLysAspPheLeuSerArgAspMetAlaMetSerGluValAlaS 361
    ||   :::       |||:::|||
114012 CAGGGATTGGCGGGCCCTTCCTCACCGATGTGGTGGCCCTCGATCTGCT 114061

361 pArgPheMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaT 378
    :  :::||::::::|||   |||:::|||
114062 GCGAGTTGGCGAT...CAGAGCGTTACGTTCAGGGCAACTCCTTGGCCA 114108

378 hrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLys 394
    |||:::|||:::|||   |||:::||||||:::
114109 CCAAGGCCATGGAGGCCATTCCCAAGCTGACGCTGGAACAGTATCTGCAG 114158

395 AspAlaIleGlyGluPheIleArgAlaLeuAlaLeuTyrGluSerGluGluAsnCy 411
    |||:::||||||:::|||
114159 GACACACTATCCGCACCAATAAACGAGCTAATTCAGTCGGAGAGGACTG 114208

411 sGluValAspProIleLys..........CysThrAlaSerSerLeuAlaG 425
    ||||||   |||:::                    ::::::||||||
114209 CGAGGTGGATCCCACCAAGACGAGCGGGTTCGTCGCTGCAGC 114258

425 luHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysVal 441
    :::||||||:::||||||   :::   :::
114259 GACACGCAGGCCGCCTTGCGTGCGGCGTCCGAGGGCGTGGCAGTGCATC 114308

442 ValAsnSerHisCysValPheProArgGluLeuLysGluValPheAlaSe 458
    |||   :::||||||:::|||   |||:::
114309 TTCGAATCGACAAGCATTTTCCCGCCCAGTTACGAAATTGCTTTGCCAC 114358

458 rTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAspArgL 475
    :  :::       |||||||||:::|||   |||:::
114359 GTTCCGGGAGCCGTTGCAGCGTGGCCCTCAGGATATGGCCCACAACC 114408

475 euIleSerAlaSerLeuArgPheLeuArgAlaIleGlnAsnLeuAsnPheSerLeuMetSer 491
    ||||||||:::   |||:::|||:::|||
114409 TGATCTCGGCGAGCATTTCCTGCGCTTCCTGCGTGCCCGGCCATCCTGTCG 114458

492 ProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnThrSerAr 508
    ||||||:::||||||||||::||:::
114459 CCGTTCCTCTTCAAATATCACCAGCCAACTGCGAACTGCGTCCGCACGGGCCACCCG 114508

508 gThrLeuThrIleLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerL 525
    |:::|||:::   :::|||   |||:::|||
114509 CAATCTCACACTGGTGCCAGACCCTGCAAACATTGGCCAATTTCACCC 114558

525 ysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGlu 541
    :::||||||:::|||:::|||   |||
114559 GCTTCCAGGCAAAGAACTTTATGGAGTTTCTCAACGATTTCCTGGAG 114608

542 LeuGluTrpGlySerMetGlnGlnPheLeuTyrGlyIleSerAsnLeu.. 557
    ||||||   :::|||||||:::||||||
```

```
 19 PheLeuSerArgArgLeuLysSerSerIleLysArgThrLysSerGlnPr 35
    ||||||   :::|||   :::::::   ::||||||||||||||
113043 TTCCTTTGCCACAGA...TCCAATCCTTTGAGCCGGACCAAGTCGGTGAC 113089

 35 oLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheArgS 52
    ||||||:::|||||:::|||
113090 CAAGCTGGACGGGACCAAG.............................. 113108

 52 erAlaAspHisAspArgAlaArgLeuMetGlnSerPheLysGluSerHis 68
       |||       |||:::          :::|||
113109 ..............CGCGGATCCGGCGACTGAGGGGCTCCCGC 113138

 69 SerHisGluSerLeuLeuSerProSerSerAlaAlaGluAla........ 82
    |||||||||||||||||    :::::  :::::
113139 TCGCACGAGAGTCTGCTGTCCACGCCGTCATCTCCACCATAGGTGA 113188

 82 ..................LeuGluLeuAsnLeuAspGluSerIleLeuPro 95
                      ||:::||:::   :::|||   |||
113189 GTTCTGGCTCATAATGGAGTTTGGCCAAACCCTCAAACCATTTTCTTT 113238

 83 ...........LeuGluLeuAsnLeuAspGluSerIleLeuPro 95
    |||:::||:::
113239 TCTTTCTTCTTCTCAGATCTCTCGACTGGGCCGTGGGCCTGGCCGCCC 113288

 96 ValHisSerSerIleLeuGlyGlnGluPheCysPheGluValThrThrSe 112
    ||||||   ||:::|||:::||:::
113289 GTGCATCAGTTCGGTTCTGGGACCTCGTCACTGTTTCCAGGTACGGGGCGG 113338

112 r...SerGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAspL 128
    |||   :::   |||:::   |||||||||
113339 GCCCGTGGCGAGCGGTACTACTCATGCGAGTCCGGATCCGCCAGGAGCCGACC 113388

128 ysTrpIleGluAsnLeuGlnArgAlaValLysProAsnLysAspAsnSer 144
    ||||||   |||:::||||||:::|||
113389 TTTGGATCTACTCGCTGCCAAGTCGATCGCTCCGAATGCAGAGACCACACG 113438

145 ArgArgValAspAsnValLeuLysLeuTrpIleIleGluLeuAlaArgGluLe 161
    |||:::||||    ||||||:::||||||:::
113439 CGTCGCACGGACAACTCGCTGAAGATGTGGGTGTACGAGGCGAAAAATCT 113488

161 uProProLysLysArgTyrTyrCysGluLeuCysLeuAspAspMetLeuT 178
    |||||||||||   :::   :::||||||||:::
113489 GCCGCCCAAGAAGCGTTACTTTTGCGAACTGCAATTGGACAAGACCTGT 113538

178 yrAlaArgThrSerLysProArgSerAlaSerGlyAspThrValPhe 194
    |:::|||||:::       :::|||
113539 ACGGCCGGACTTCGATTCCGGCAGTGAAGCTGCAGACG........GATCTGCTGTTT 113579

195 TrpGlyGluHisPheGluAsnAsnLeuProAlaValArgAlaLeuAr 211
    ||||||:::||||||:::|||:::|||
113580 TGGGGAGAGCACTTCGAATTCCCAATATCACCAGCCGAGATTAATGTGATCAC 113629

211 gLeuHisLeuTyrArgAspSerAspLysLysArgLysLysAspLysAlaG 228
    |||:::|||||||:::||:::
113630 TGTTAACGTTTTCCGTGAGGTGGACAAGAAGAAGCGGGACAAATACC 113679

228 lyTyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgHis 244
    |:::||||   ::::::|||||||:::
113680 AATTTGTGGGATCGGTGAAGATACCCGTCACGGATGTCACCTCCAGATTG 113729
```

Fri Jan 19 06:20:05 2001

us-09-294-298-6.rge

```
115332  CTT.........                                                             115334
                   |||::::                      ||||||||||
807     nLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaA   824

115334  .........                                                               115334
                                               |||
824     spGlyProGlyProAlaGlyHisGlyGlySerGly......His          838
                                            .....AATGGGAGCAGCAGCAATTTGCTGCAC      115361
115335  .........                                                               115361

839     GlyProProSerHisHisHisHisHisHis..........               Hi    851
            |||       |||::::::|||   ||||||
115362  GGC.......   CACCAGCAGCATGCCCATCACCCGCAGCAACTGCA         115399

851     sHisHisArgGlyGlyGlyGluProProGlyAspThrPheAlaProPheHisG   868
           |||::::|         ||||||
115400  TCCACATCACTGC......  CCGCCG.                             115418

868     lyTyrSerLysSerGluAspLeuSerThrGlyValProLysProProAla     884
                                          ::::::|   |||
                              GCGCCACAGACCAGTGCC                  115436
115419  .........                                                               115436

885     AlaSerIleLeuHis..........               SerHisSerTyrSerAspGl  896
          :::|||  ::::::|                       :::::|  |||   :::::|
115437  TCCAGCACTATGGAGCGCCATGGATCGCATGAACTATCCGTATATGTCGCA          115486

896     uPheGlyProSerGlyThrAspPheThrArgArgGlnLeuSerLeuGlnA     913
                     :                  |||
115487  C.......    AATGGCAATGACTACCAG                          115505

913     spAsnLeuGlnHisMetLeuSerProProGlnIleThrIleGlyLyProGln   929
                          |||          |||   :::
115506  .........    ACCAGCACGCCCTTCGACGCACTCGCTCCAGGACA        115538

930     ArgProAlaProSerGlyPro.......   GlyGlyGlySerGlyGlyGlySe  944
                                       :::::|||::|:|
115539  CTGCCCACGGAATGGAAATCCCAATGCCAATGGCAACGTGGGCAGCAGCAA        115588

944     rGlyGlyGlyGlyGlyGlnProProPro.LeuGlnArgGlyLysSer        960
                                       :::|||||
115589  CAATAACCAGAGCGGCAGCTACGATGACATGCACGGGGAGTTCCAAATCC        115638

961     GlnGlnLeu.......                  ThrValSerAlaAlaGlnLysProAr  972
               :::|||                      ::::::|||:::|::|::
115639  AGATCTCTGGGGTTCGATACGAGCACGTGCTTTGTCTGCAAGTCGCCCACA       115688

972     gPro.......         Ser. SerGlyAsnLeuLeuGlnSerPro      982
          :::|                  ||:::|||:
115689  CCCATGATGAAATCCAGTTTGGGACCAGCCGGGCCCGACGAGGTCATCA        115738

983     GluProSerTyrGlyProAlaArgProArg..........               992
          ::  ::::::|||  ::::|||
115739  CAAACTGAATTTGGGAATACCCGATCACCTCAGGTGGCTATGTCGGGGGTA        115788

993     .........    GlnGlnSerLeuSerLysSerGluGlyGlySerIleGlyGlySerGlyG  1007
                     ::::::|||:::::||||::::|||::||||
115789  ATAATTTGAATCCCAACTCGAATATGCCAAGAACTTGGAGGATCTGGAC        115838

1007    lySer..........               GlyGlyGlyGlyGlyGlyLeuLysPro  1018
          :::|                         ||||||      ::::|||
115839  GATCTGTTCAAGTACCGCCGGGAGCGCATGAGCTGGCGGAACCAGCAACCA        115888

1019    SerIleThrLysGlnHisSer...GlnThrProSerThr.LeuAsnPro.     1033
          :::|||||       |||:::  |||
115889  TCACAATCACACCAGGGTCAGCAGAACCACCAGGGACATCTGAGGCCGG        115938

1034    ..ThrMetProAlaSerGluArgThrValAlaTrpValSerAsnMetPro     1049
                |||         |||
115939  CCGCCGTTCCCGGCAAGGAGCAGCTGTCGCGGAAAAGCAGT              115979

1050    HisLeuSerAlaAspIleGluSerAlaHisIleGluArgGluTyrLy        1066
          |||   |||    ::::::|||
115980  CACTGCCAGTTCTGGCTACCAGAGCATCTCCACAAATCCC.                   116018
```

```
114609  CAGGAGGCCGCTCCCATGCAACAGTTTCTGGAGATTATATCCACGCGGCC    114658
            |||     ::::                  ||||||||||||
                                          |||||||||||
558     .........    AspThrLeuThrAsnSerSerPheGluGlyT   569
            |||:::::|          ::    ||||
114659  GGAGCACCCAGCTCCAGACTCCGATCCTCGAT.......TGGGCCGGGT    114699

569     yrIleAspLeuArgGluLeuSerThrLeuHisAlaLeuLeuTrpGlu     585
          ||||||||  |||:::::||   |||::|
114700  ACATCGACCACCAGGGCAAACAGTTGTCCATCTACACACAGTTGCTCAGCGAA  114749

586     ValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGlyLeuProLeuPr  602
          |||  ::::||       :::||   ||||||
114750  AGCCTGGCCAAGCTGCCGGAGGCCAGCAGGCCAGCGAGCTGGATCCGTTGCA   114799

602     oArgLeuLeuSerAspIleSer..........               ThrA    611
          |||:::::|||::|::|:|                           |||
114800  GCATATTCTCGATGAAATCAGCCGAGCCCAAAGAGCATGGCATGGGCACAG   114849

611     laLeu.........                                        612
          |||||
114850  CACTGCCGCGGTGGATATTTGCCGCCCACCTGCTCTACGCACTCGATAGCC   114899

613     .........    ArgAsnPro..........               615
                       ||| ||
114900  AGCGAGAATCAGGAGAATCGCAATCCGGGATCCTCGGCGTCTCCCACCCTGG   114949

616     ...AsnIleGlnArgGlnProSerArgGlnSerGluArgGlaArgSerG    631
            ::::||||       ||||||||::::    ||||::
114950  CTCCAACTCGGAGCAGTTACTGCCACAACAAAGCCAGTTGGCCCAGCCGC    114999

631     lnProMetValLeuArgGlyPro...SerAlaGlu..........     MetGln  643
          |||       ||||||||||   ||:::||||||
115000  AGCATGGGATTGTTAGTAAACCATTATCTGCGGAGCCGGCCGCATCATGCCA   115049

644     Gly.TyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerProMet  660
          |||  |||||||:::||   |||:::|||
115050  GGAGTACTTACGCCGCGAATTCTCTGGAGAAGATATCTTTAGATACAATGA   115099

660     etAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerPro   676
          |||::::||       ||||||   ||||||
115100  TCCCACGGTTAATG.......                                115113

677     ThrLysGluLysProProProProProGlyGlyGlyLysAspLeuPh     693
          :::|||
115113  .........                                           115113

693     eTyrValSerArgProProLeuAlaArgSerSerProAlaTyrCysThrS   710
                                          ||||||||||||:::|
115114  GCTTACTGCAGCA                                       115126

710     erSerSerAspIleThrGluProGluGlnLysMetLeuSerValAsnLys   726
          |||||||||||      |||    |||
115127  GCAGCAGCAGCAGC.......                               115158

727     SerValSerMetLeuAspLeuGlnGlyAspLeuGlyProGlyLyArgLeuAs  743
          |||   |||      :::||
115159  AGCAACAGCAACATC.......                              115184

743     nSerSerValSerAsn.......  LeuAlaAlaValGly..AspLeuLe    757
          ||||       ::::::||||      ::::::||
115185  CAGCATGGCCATCGCAACAGCCCACCACCAGCCATCCACTCCAGATGCT   115234

757     uHisSerSerGlnAlaSerLeuThrAlaLeuAlaLeuGlyAlaArgProAlaP  774
          |   :::::::::|||   |||::   :::||||
115235  CTCCAATTCACAAACCTCCATT..GCCGGCAACCAATATATGAGTTCGC   115281

774     roAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAlaGly   790
          ||:::||   :::|||   :::||||::::::::::
115282  CAGGAGGCCTGCAGCATCGCAGACCTCGATGGCCGTCCCATCG         115331

791     MetArgLeuSerGlnMetGlyValThrArgAspGlyValProAlaGlnGl   807
          :::
```

Fri Jan 19 06:20:05 2001

us-09-294-298-6.rge

```
*            1       509:  contig of 509 bp in length
*          510       589:  gap of unknown length
*          590      1201:  contig of 612 bp in length
*         1202      1281:  gap of unknown length
*         1282      2394:  contig of 1113 bp in length
*         2395      2474:  gap of unknown length
*         2475      3512:  contig of 1038 bp in length
*         3513      3592:  gap of unknown length
*         3593      4500:  contig of 908 bp in length
*         4501      4580:  gap of unknown length
*         4581      5465:  contig of 885 bp in length
*         5466      5545:  gap of unknown length
*         5546      6253:  contig of 708 bp in length
*         6254      6333:  gap of unknown length
*         6334      7105:  contig of 772 bp in length
*         7106      7185:  gap of unknown length
*         7186      8077:  contig of 892 bp in length
*         8078      8157:  gap of unknown length
*         8158      9047:  contig of 890 bp in length
*         9048      9127:  gap of unknown length
*         9128      9530:  contig of 403 bp in length
*         9531      9610:  gap of unknown length
*         9611     10173:  contig of 563 bp in length
*        10174     10253:  gap of unknown length
*        10254     10841:  contig of 588 bp in length
*        10842     10921:  gap of unknown length
*        10922     11807:  contig of 886 bp in length
*        11808     11887:  gap of unknown length
```

```
1066 sLeuLysGluTyrSerLysSerMetAspGluSerArgLeuAspArgValL 1083
116019 ................TCGCCCTGCGAGTCCTCCAGTCCCGTGGAG....... 116048
1083 ysGluTyrGluGluIleHisSerLeuLysGluLysGluArgLeuHisMetSer 1099
                   ::|!!!!|        ::|:
116049 .................AGCCAGCTGAAGGCCGCGATGGGCAGTCAC 116078
1100 AsnArgLysLeuGluTyrGluArgGluLeuSerGlnGlnGluGluGluGlul 1116
     !!!     |!!!||       ::|  |!!||||
116079 AATGCGACGCCGCTG...GCCTTCAAGAATCCCTCCTATCAGCTTCAGCCCCA 116125
1116 nThrSerLysIleLeuMetGlnTyrGlnAlaArgLeuGluGlnSerGluL 1133
     !!!!::|::      ||  |!!|:::::::
116126 AACTGGCTTCGTCCAGATCATCGGCACAGAGTAATACACACCAGCAGCAGC 116175
1133 ysArgLeuArgGlnGlnValGluLysAspSerSerGlnIleLeuLysSerIle 1149
     ::|:   :|||||||||         |||||:::
116176 AGCAGCAGCAACAACAGTT......GGCAGCCGCCTTGAAACCAATT 116219
1150 IleGlyArgLeuMetLeuValGluGluGluLeuLeuArgArgAspHisProAl 1166
     |||   |!!|: :::::::: |!!||||||||||||        |!
116220 GGCGGCGGACTGGTGCCCGCGAGGGCGGCTTTCCTCAACAGTGGCGGAGC 116269
1166 aMet.AlaGluProLeu........ProGluProLys............. 1175
     |:::  ||:|||||:| |||  !!!
116270 CTTGGAGGCGGCCACTTTGACCCCCAGCTCCTCGGACGAACAGCTCTCGG 116319
1176 ................LysArgLeuLeuAspAlaGlnIleArgGlySe 1185
                    :::|!!!|        ::|!!!|
116320 CGGATAATTACTTCAGTTATGCAGCGGCTGCAGCTGCTGAGCGGGTATT 116369
1185 rPheProProTrpValGlnInThrArgValArgTrpProHisLeuGlyT 1202
     |||   !!!     |!!|::        :::|
116370 GCCACCAAATTGGAGGCTCAACGCTCGC........TCAGCGCGGCGCAG 116410
1202 hrAlaTrpProProGln...................... 1207
     ::|!!|   !||||||||||
116411 TAGCTCCTCCAACCTCAGCATCTGCGTCCACCTCGAATCTTGGCAAGAGCG 116460
1208 ...ProHisProProGlyCysArgSerGlnArgThrAlaSerSe 1223
     !||||||||   |||    :::|!!|        ::|!!||
116461 GCGGCTCATCCGCCTACGGGCGCGCTGAATGGGCCGCTGAAGCGCGAGGAT 116510
1223 rGlyThrProGlnInThrSerProPro............SerIleThrA 1236
     :::|!!|||||:
116511 GTCTACGCGCAGTGGCTACGGCCGCAGCAGTGGAAATGTGGGCTATGGCTT 116560
1236 spLeuLeuPro.Cys.........AlaProTyrProGlyProSerValTh 1250
     |!!|||| ||        ::|!!||||||:::::::::
116561 GTCCACTTCCAGTGCCGCGGGACACCATCAACATCCCCACCAGCAGCAGC 116610
1250 rAspLeuLeuProSerAlaProAlaProAspProGlyThrSerProThrThrArgT 1267
     ::|!!|        |!!!:::::
116611 AGAATCCGATGCAGCAGCAGCAGCAAAGGGAACGGGATCAGGAACACAAG 116660
1267 hr............GlyArgHisGlnGlyThrAlaGlyLeuSerPro 1278
     ::         |||   ::|!!| ::|!!|||
116661 CAGTATGCCGGGCAGTGTGCCGGGCACCGTGGGATCGGCCACATCAGCGGC 116710
1279 LeuThrProPro........TrpGlyThrHis 1286
     |||   !!!|||:::|!!|
116711 TCAGAGGGCCCTGAGCTTGGACTCGGCCGCCAC 116743
```

seq_name: gb_htg4:AC012161

```
* 11888  13165: contig of 1278 bp in length
* 13166  13245: gap of unknown length
* 13246  14526: contig of 1281 bp in length
* 14527  14606: gap of unknown length
* 14607  15801: contig of 1195 bp in length
* 15802  15881: gap of unknown length
* 15882  16640: contig of 759 bp in length
* 16641  16720: gap of unknown length
* 16721  17679: contig of 959 bp in length
* 17680  17759: gap of unknown length
* 17760  18797: contig of 1038 bp in length
* 18798  18877: gap of unknown length
* 18878  19955: contig of 1078 bp in length
* 19956  20035: gap of unknown length
* 20036  20931: contig of 896 bp in length
* 20932  21011: gap of unknown length
* 21012  21643: contig of 632 bp in length
* 21644  21723: gap of unknown length
* 21724  23416: contig of 1693 bp in length
* 23417  23496: gap of unknown length
* 23497  24457: contig of 961 bp in length
* 24458  24537: gap of unknown length
* 24538  25955: contig of 1418 bp in length
* 25956  26035: gap of unknown length
* 26036  27838: contig of 1803 bp in length
* 27839  27918: gap of unknown length
* 27919  29686: contig of 1768 bp in length
* 29687  29766: gap of unknown length
* 29767  30976: contig of 1210 bp in length
* 30977  31056: gap of unknown length
* 31057  32460: contig of 1404 bp in length
* 32461  32540: gap of unknown length
* 32541  34644: contig of 2104 bp in length
* 34645  34724: gap of unknown length
* 34725  36341: contig of 1617 bp in length
* 36342  36421: gap of unknown length
* 36422  37864: contig of 1443 bp in length
* 37865  37944: gap of unknown length
* 37945  39424: contig of 1480 bp in length
* 39425  39504: gap of unknown length
* 39505  41341: contig of 1837 bp in length
* 41342  41421: gap of unknown length
* 41422  42970: contig of 1549 bp in length
* 42971  43050: gap of unknown length
* 43051  44419: contig of 1369 bp in length
* 44420  44499: gap of unknown length
* 44500  46484: contig of 1985 bp in length
* 46485  46564: gap of unknown length
* 46565  48446: contig of 1882 bp in length
* 48447  48526: gap of unknown length
* 48527  50693: contig of 2167 bp in length
* 50694  50773: gap of unknown length
* 50774  53578: contig of 2805 bp in length
* 53579  53658: gap of unknown length
* 53659  55753: contig of 2095 bp in length
* 55754  55833: gap of unknown length
* 55834  58457: contig of 2624 bp in length
* 58458  58537: gap of unknown length
* 58538  60171: contig of 1634 bp in length
* 60172  60251: gap of unknown length
* 60252  63621: contig of 3370 bp in length
* 63622  63701: gap of unknown length
* 63702  66219: contig of 2518 bp in length
* 66220  66299: gap of unknown length
* 66300  69074: contig of 2775 bp in length
* 69075  69154: gap of unknown length
* 69155  72905: contig of 3751 bp in length
* 72906  72985: gap of unknown length
* 72986  75963: contig of 2978 bp in length
* 75964  76043: gap of unknown length
* 76044  78866: contig of 2823 bp in length
* 78867  78946: gap of unknown length
* 78947  82045: contig of 3099 bp in length

* 82046   82125: gap of unknown length
* 82126   85135: contig of 3010 bp in length
* 85136   85215: gap of unknown length
* 85216   88550: contig of 3335 bp in length
* 88551   88630: gap of unknown length
* 88631   92611: contig of 3981 bp in length
* 92612   92691: gap of unknown length
* 92692   98709: contig of 6018 bp in length
* 98710   98789: gap of unknown length
* 98790   104147: contig of 5358 bp in length
* 104148  104227: gap of unknown length
* 104228  111796: contig of 7569 bp in length
* 111797  111876: gap of unknown length
* 111877  118488: contig of 6612 bp in length
* 118489  118568: gap of unknown length
* 118569  126924: contig of 8356 bp in length
* 126925  127004: gap of unknown length
* 127005  137051: contig of 10047 bp in length
* 137052  137131: gap of unknown length
* 137132  148859: contig of 11728 bp in length
* 148860  148939: gap of unknown length
* 148940  149554: contig of 615 bp in length
* 149555  149634: gap of unknown length
* 149635  150197: contig of 563 bp in length
* 150198  150277: gap of unknown length
* 150278  150907: contig of 630 bp in length
* 150908  150987: gap of unknown length
* 150988  151646: contig of 659 bp in length
* 151647  151726: gap of unknown length
* 151727  152074: contig of 348 bp in length
* 152075  152154: gap of unknown length
* 152155  152902: contig of 748 bp in length
* 152903  152982: gap of unknown length
* 152983  153624: contig of 642 bp in length
* 153625  153704: gap of unknown length
* 153705  154242: contig of 538 bp in length
* 154243  154322: gap of unknown length
* 154323  154739: contig of 417 bp in length
* 154740  154819: gap of unknown length
* 154820  155338: contig of 519 bp in length
* 155339  155418: gap of unknown length
* 155419  155993: contig of 575 bp in length
* 155994  156073: gap of unknown length
* 156074  156806: contig of 733 bp in length
* 156807  156886: gap of unknown length
* 156887  157541: contig of 655 bp in length
* 157542  157621: gap of unknown length
* 157622  158208: contig of 587 bp in length
* 158209  158288: gap of unknown length
```

```
alignment_scores:
    Quality:  1257.50                    Length:   1432
    Ratio:    1.642                      Gaps:     57
    Percent Similarity: 53.492   Percent Identity: 29.190

alignment_block:
US-09-294-298-6 x AC012161/rev  .

Align seg 1/1 to reverse of: AC012161  from: 1  to: 202741

    19 PheLeuSerArgLeuLysSerIleLysSerThrLysSerGlnPr 35
       ||||||      ::::::    ::||||||||||||||||
126785 TTCCTTGCCACAGA...TCCAATCCTTGAAGCGGACCAAGTCGGTGAC 126739

    35 oLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheArgS 52
       |||||::||||||:::::::
126738 CAAGCTGGAGCCGGACCAAG.................. 126720

    52 erAlaAspHisAspArgAlaArgLeuMetGlnSerPheLysGluSerHis 68
                              ::::::: |||     |||:::
126719 .....................CGCGGATCCGGCGGACTGAGGGGCTCCCGC 126690

    69 SerHisGluSerLeuLeuSerProSerSerAlaAlaGluAla......... 82
```

```
125816 CAGGGATTGGGGGGCGCCTTCCTCACCGATGTGGTGGCCCCTCGATCTGCT 125767
   361 pArgPheMetGluArgGluHisLeuIleLeuIlePheArgGluAsnThrLeuAlaT 378
       |||      :::       ||||||  ||||||||  ||||||||  ||||||||
125766 GCGAGTTGGCGAT...CAGAGGCTTACGTTCAGGGGCAACTCCTTGGCCA 125720

   378 hrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLys 394
       ||||:||:::|||    ||||||:::|||   ||||||||  ||||||||
125719 CCAAGAGCATGGAGCATTCCTCAAGCTGACGGGCGAACAGTATCTGCAG 125670

   395 AspAlaIleGlyGluPheIleGluArgAlaLeuTyrGluSerGluGluAsnCy 411
       ||:::::::::            |||::  |||    :: ]|||||::::::]
125669 GACACACTATCCGCACCAATAAACAGCTRATTCAGTCGAGAGGGACTG 125620

   411 sGluValAspProIleLys..........CysThrAlaSerSerLeuAlaG 425
       ||||||||||  |||         ::||||    :::||||||:|||
125619 CGAGGTGGATCCCACCAGACGAGCGGTTCGTCGGCGGGTTCGCTCAGC 125570

   425 luHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysVal 441
       :::|||||||  ||||||||||||   ||:::|||||||||||||||||||
125569 GACAGCAGGCCCCCTTGCGTGGCGCGTCCAGGGGCGTGGCAGTGCATC 125520

   442 ValAsnSerHisCysValPheProArgGluLeuLysGluValPheAlaSe 458
       ::||||||]  ||||||    ||||||::||||::  :: ||||||
125519 TTCGAATGCCACAAGCATTTTCCGCCCAGTTACGAAAATTGCTTTGGAC 125470

   458 rTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAspArgL 475
       ::||||||  |||    ||||||::  ::  |||||||||| |||::::::
125469 GTTCCGGGAGCGTTGCAGCAGTCTGGGCCGTCAGGATATGGCCGACAACC 125420

   475 euIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSer 491
       ||||||||||||||  |||||||||||:::|||  ||:::||||:::|||
125419 TGATCTCGGCGAGCATTTTCCTGCCGCTTCCTGTGCCGGCCATCCTGTCG 125370

   492 ProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnThrSerAr 508
       :::||||||  ||||||  |||::  |||    ::::||||||  |||||
125369 CCGTCGCTCTTCAATATCACCAGCGAACTGCCGTCCGACACGGGCCACCCG 125320

   508 gThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerL 525
       ||:::|||||||:|||||::  ||||||  |||    ||||||  ||||||
125319 CAATCCACACTGGTGCCAAGACCCTGCAAACATTGGCCAATTTCACCC 125270

   525 ysPheThrSerLysGluAspPheMetAsnGluPheLeuGlu 541
       ||||||    ::||||||  |||::  ||||||::||||||
125269 GCTTCCAGGCAAAGAGAACTTTATGGAGTTTCTCAACGATTTCCTCGAG 125220

   542 LeuGluTrpGlySerMetGlnGlnPheLeuTyrGluIleSerAsnLeu.. 557
       |||   |||::||||||    |||    |||||||::||||||||  ::
125219 CAGGAGCCGGTCGCATGCAACAGTTTCTGGAGATTATATCCAGCGGCC 125170

   558 ..........AspThrLeuThrAsnSerSerSerPheGluGlyT 569
                :::       :::       ||  :::      ||||
125169 GGAGCACCCAGCTCCAGACTCGAT......TGGGCCGGGT 125129

   569 yrIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuTrpGlu 585
       |:::|||  ||||||||:::||||::||||  ::::||||||||  ||
125128 ACATCGACCAGGGCAAACAGTTGTCCATACTACACAGTTTGCTCAGCGAA 125079

   586 ValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGlyProLeuPr 602
       ::||||||  ||||::  ::::|||    ||:::|||||||
125078 AGCCTGGCCAAGCTGCCGGAGGCCAGGCCACGAGCTGGATCCGTTGCA 125029

   602 oArgLeuLeuSerAspIleSer..............ThrA 611
       ::||||||||||||||:::  ::::         ||||
125028 GCATATTCTCGATGAAATCAGCCGAGCCAAAGAGCATGGCACACAG 124979

   611 laLeu.......................ArgAsnPro........... 615
       ||||
124978 CACTGCCGGGTGGATATTTGCCGCCACCTCGTCTACGCACTCGATAGCC 124929

   613 ..............ArgAsnPro...........
124928 AGCCAGAATCAGGAGAATCGGAATCCGGGATCCTCGGGCTCCACGCTGG 124879
```

```
126689                     ||||||||||||||||            ::::::
       TCGCACGAGAGTCGTGCTGTCCAGTCACGCCGTCATGTCCACCATAGGTGA 126640

    82 ...................                                 82
                                                    126639
       GTTCTGGCTCATAATGGAGTTTGGCCAAACCCTCAAACCATTTTCTTTT 126590

    83 ............LeuGluLeuAsnLeuAspGluAspSerIleIleLeuLysPro 95
                    ]|:|:::|||   ||||||||    ::::|||
126589 TCTTTCTTCTTCTAGATCTCTCGTGCACTGGGGCCGGTGGGCGTGGCGCCC 126540

    96 ValHisSerIleLeuGlyGlnGluPheCysPheGluValThrThrSe 112
       ||||||::::|||::  ||||||||||  ||||||  :: ::|||||
126539 GTGCATCAGTCGTTCTGGGACGTCGTCACTGTTTCCAGGTACGGGCGG 126490

   112 r...SerGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAspL 128
       :    :: ::|||||||||    ||||||:::|||    ||||||
126489 GCCTCGGTGCGAGCCGGTACTACTCATGCGGATACGCCCAGGAGCCGCACC 126440

   128 ysTrpIleGluAsnLeuGlnArgAlaValLysProAsnLysAspAsnSer 144
       ||||||  :::||||||||||||    ||||||::::::::
126439 TTTGGATCTACTCGCTGCCAAGTCGATCCGCTCGACCAGCACACG 126390

   145 ArgArgValAspAsnValLeuGlyLeuTrpIleIleGluAlaArgGluLe 161
       ||||||  |||::||:::|||||  ::||||||::||||||||  ::||
126389 CGTCGCACGGACAACTCGCTGAAGATGTGGGTGTACGAGGCGCAAAAATCT 126340

   161 uProProLysArgTyrTyrCysGluLeuCysLeuAspAspMetLeuT 178
       ||||||||||||||||    :::|||   |||::|||||||  |||
126339 GCCGCCCAAGAGGCGTTACTTTTGCGAACTGCAATTGGACAAGAGACGCTGT 126290

   178 yrAlaArgThrThrSerLysProArgSerAlaSerGlyAspThrValPhe 194
       ||||||  ||||||  ::||||||||||||||    ::  |||
126289 ACGGCCGGACTTCGGTGAAGCTGCAGACG.......GATCTGCTGTTT 126249

   195 TrpGlyGluHisPheGluHisAlaAsnLeuProAlaValArgAlaLeuAr 211
       |||||||||::||  :::|||    :::|||    ::|||
126248 TGGGGGGAGCACTTCGATTTCCCCGACATACCCGAGATTAATGTGATCAC 126199

   211 gLeuHisLeuTyrArgAspSerAspLysLysArgLysLysAspLysAlaG 228
       |:::|||||||:::||||||||   |||||||::||||||||||||||
126198 TGTTAACGTTTCCGTGAGGTGGCAAGAAGAAGAGCGGGACAAATACC 126149

   228 lyTyrValGlyLeuValAlaThrValProValAlaThrLeuAlaGlyArgHis 244
       ::::|||||    |||    ::|||||||||||:::    |||:::|||
126148 AATTTGTGGATCGGTGAAGATACCCGTGCACGATGTCACCTCCAGATTG 126099

   245 PheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGlySe 261
       |||||||||||:::|||    ::::||| |||  :||||||||||||
126098 CCCTGGCAACAATGGTAATCCCCATA...CTGAGCGACAAGGCGGCGACAG 126052

   261 rGlyGlyMetGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyG 278
       ||    :::|||||||||||||||||||||||||||||||||
126051 TCTGGGCAGGAGACATCGGGC.............GGCGGCGCAGTGGGT 126017

   278 lyLysGlyLysGlyGlyGlyCysProAlaValAlaSerAlaLeuValAlaArgLeuLysSerAlaArgTyrGln 294
       ::||||||  |||   ::  ||||||::|||:::|||||  ||||||:::||
126016 CCAAGGACAAGGAGCAATTGCCCACCGCTGAGGATCAAGTGTCGTTTCCAG 125967

   295 ThrMetSerIleLeuProMetGluTyrLysGluGluProAlaGluTyrTyrVa 311
       ||    ::||||||||||||  ::|||||||||||||||||:::||||||||
125966 AGCACCGACACTCCTGCCCATCAATGTGTACGGACAACTTTTTGACGTACCT 125917

   311 lThrAsnHisTyrArgMetGlyCysAlaValLeuGluCysAlaLeuGluAsnV 328
       :::::::::|||::::  ::|||||||||||||||||||||:::::::|
125916 CAAGGAGAACTATAAGCGGTGCGAGACCCCTGAGACCCCGGTGATCGGAG 125867

   328 alLysGlyLysGlyGlyCysIleValAlaSerAlaLeuValAlaIleHisIleLeuGlnSer 344
       ::||||||    |||||||||||  ||||||:::|||  :::||:::|||
125866 TCAAGGGCCAAGGAGGACATTGGACAGCCACTGGTGCTCATGCACGCA 125817

   345 ThrGlyLysAlaLysAspPheLeuAsnSerAspMetAlaMetSerGluAlaSer 361
       |||    |||       ::::::::::::   :::::::
```

us-09-294-298-6.rge

Fri Jan 19 06:20:05 2001

```
896  uPheGlyProSerGlyThrAspPheThrArgArgGlnLeuSerLeuGlnA  913
         ::::|||::|||::|::
124341 C........AATGGCAATGACTACGAG.......                    124323

913  sProAsnLeuGlnHisMetLeuSerProProGlnIleThrIleGlyProGln  929
                        |||||
124322 ........ACCAGCACGCCTTCGACCACTCGCTTCCAGGACA           124290

930  ArgProAlaProSerGlyPro......GlyGlyGlySerGlyGlyGlySe   944
     |||               ::|||||         |||::|||||
124289 CTGCCACGGAATGGAAATCCCAATGGCCAATGGGCACCAGCAGCAA       124240

944  rGlyGlyGlyGlyGlyGlyGlnProProPro.LeuGlnArgGlyLysSer   960
                                     ::|||||
124239 CAATAACCAGAGCGGCAGCTACGATGACATGACACGGGGAGTTCCAAATCC  124190

961  GlnGlnLeu...........ThrValSerAlaAlaGlnLysProAr      972
     :::::||||          :::::||::|||::::|||::
124189 AGATCTCTGGGTTCGATACGAGCAGTGCTTTTGTCTGCAAGTCGCCCACA  124140

972  gPro........................Ser.SerGlyAsnLeuGlnSerPro  982
     :|                          ::|||::
124139 CCCATGATGAAATCCAGTTTGGGACCAGCCGGGGGGACGAAGTCATCA     124090

983  GluProSerTyrGlyProAlaArgProArg...............       992
     :::::|||:|  ::|||
124089 CAAACTGAATTTGGGAATACCCGATCACTCAGGTGGCTATGTCGGGGTA   124040

993  ......GlnGlnSerLeuSerLysGluGlySerIleGlyGlySerGlyG   1007
           ::::|||::|||       :::::||||||||||||
124039 ATAATTTGAATCCCAACTGGAATATGCCCAAGGACTTGGAGGATCTGGAC  123990

1007 lySer...........GlyGlyGlyGlyGlyGlyGlyGlyLeuLysPro   1018
        |||         |||   ||::|||||||||||
123989 GATCTGTTCAAGTACGCCGAGGAGCATGACGTGGCGGAACCAGCGAACCA  123940

1019 SerIleThrLysGlnHisSer...GlnThrProSerThr.LeuAsnPro.  1033
     |||  ::::::|           |||::|||  |||::|||
123939 TCACAATCACAACCAGGGTCAGCAGAACCACCCAGGGACATCTGAAGCCGG 123890

1034 ..ThrMetProAlaSerGluArgThrValAlaTrpValSerAsnMetPro  1049
                 ::::|||            |||||    |||
123889 CCGCCGTTCCCGGCAAGGAGCAGCTGTCGGCGAAAAGCAGT.........  123849

1050 HisLeuSerAlaAspIleGluSerAlaHisIleGluArgGluGluTyrLy  1066
     |||  ::|||        ::|||
123848 CACTGCAGTTCTGGCTACCAGCAATCTCCACAAATCCC............  123810

1066 sLeuLysGluTyrSerLysSerMetAspGluSerArgLeuAspArgValL  1083
                        ::|||          |||
123809 .............TCGCCCTCGCAGTCCTCCAGTCCGTGGAG.......  123780

1083 ysGluTyrGluGluIleHisSerLeuLysGluArgLeuHisMetSer     1099
                                ::|||||
123779 ..........AGCCAGCTGAGGGCCGATGGGCAGTCAC              123750

1100 AsnArgLysLeuGluTyrGluArgArgLeuLeuSerGlnGluGluGluGl  1116
     |||                       :::  :|||
123749 AATGCGGCCGCTG...GCCTTCAAGAATCCCTCCTATCAGCTTCAGCCCCA 123703

1116 nThrSerLysIleLeuMetGlnTyrGlnAlaArgLeuGluGlnSerGluL  1133
     ::|||                   |||          |||
123702 AACTGGCTCGTCCAGATCATCGGCACAGAGTAATACACCAGCAGCAGAGC  123653

1133 ysArgLeuArgGlnGlnGlnValGlnLysAspSerGlnIleLysSerIle  1149
     :::::  ::|||||||        |||     |||
123652 AGCAGCAGCAACAACAACAGTT......GGCAGCCGCCGTTGAAACCAATT 123609

1150 IleGlyArgLeuMetLeuValGluGluGlnLeuArgArgAspHisProAl  1166
     ::|||        ::|||       |||
123608 GGCGGCGGACCTGGTGGCCGCGAGGGCGGCCTTTCCTCAACAGTGCGGAGC 123559

1166 aMet.AlaGluProLeu.......ProGluProLys.               1175
```

```
616  ...AsnIleGlnArgGlnProArgGlnSerGluArgAlaArgSerG      631
        :::|||::|             |||
124878 CTCCAACTCGGAGCAGTTACTGCCACAACAAGCCAGTTGGCCCAGCCGC   124829

631  lnProMetValLeuArgGlyPro...SerAlaGlu........MetGln  643
     |::|                      |||||||            |||::
124828 AGCATGGCGATTGTTAGTAAACCATTATCTGCGGAGCGGCCGCATCATGCGA 124779

644  Gly.TyrMetArgAspLeuAsnSerSerIleAspLeuGlnInSerPheM  660
     |||  |||::||   |||::||||||||::|||
124778 GGAGTACTTACGCCGAATTCTCTGGAGAAGAATATCTTTAGATACAATGA  124729

660  etAlaAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerPro 676
     ::  |||  :::|||
124728 TCCCACGGTTAATG.......................                124715

677  ThrLysGluLysProProProProProGlyGlyGlyLysAspLeuPh    693
124715 ....................................              124715

693  eTyrValSerArgProProLeuAlaArgSerProAlaTyrCysThrs   710
     :::|||            ||||||
124714 ...................GCTTACTGCAGCA                    124702

710  erSerSerAspIleThrGluProGlnLysMetLeuSerValAsnLys    726
     ||||||::::  ::::::|           |||  :::::||
124701 GCAGCAGCAGCAGCAGC.........AGCAGCAACAAC               124670

727  SerValSerMetLeuAspLeuGlnGlyAspGlyProGlyGlyArgLeuAs 743
     |||  |||                        |||
124669 AGCAACAGCAACATC.....................AGCAGCTGCAA     124644

743  nSerSerSerValSerAsn.....LeuAlaAlaValGly..AspLeuLe  757
     ||||  ::::::||||||      :::::||||           ::|||
124643 CAGCATGGCCATCAACAGCCACCAGCCACCAGCCATCCCAGATGCT      124594

757  uHisSerGlnAlaSerLeuThrAlaLeuLeuGlyLeuArgProAlaP    774
     :::|||||||::|||::||       |||                  ::|
124593 CTCCAATTCACAAACCTCCATT...GCCGCCAACCAATATGAGTTCGC    124547

774  roAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAlaGly  790
     |::|||||||      ::::::|||       |||
124546 CAGGAGGCCTGCAGCATGCCAATCGCAGACCTCGATGGGCGTCCTCATCG  124497

791  MetArgLeuSerGlnMetGlyValThrThrArgSpGlyValProAlaGlnIl 807
     |||      |||       :::::|||
124496 CTT......                                           124494

807  nLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaA  824
     :::|||
124494 ...................................               124494

824  spGlyProGlyProAlaGlyHisGlyGlySerGly......His      838
     |::|||:::|||||||                          |||
124493 ..............AATGGGAGCAGCAGCAATTTGCTGCAC           124467

839  GlyProProSerHisHisHisHisHis................H1     851
     |||::::::|||||||||  ||||||                 ||
124466 GGC..........CACCAGCAGCATGCCCATCACCCCGCAGCAACTGCA   124429

851  sHisHisArgGlyGlyGluProProGlyAspThrPheAlaProPheHisG  868
     |  :::|||       ||||||
124428 TCCCACATCACTGC.......CCGCCG...                      124410

868  lyTyrSerLysSerGluAspLeuSerThrGlyValProLysProProAla 884
                        ::::::|||||||||
124409 ..................GCGCCACAGACCAGTGCC                124392

885  AlaSerIleLeuHis................SerHisSerTyrSerAspGl  896
     ::|||||               |||  :::|||||
124391 TCCAGCACTATGGAGCCGACATGGATCGCATGAACTATCCGTATATGTCGCA 124342
```

```
123558  |:::  :::|||||      |||     |||:::
        CTTGGAGGCGGCCACTTTGACCGCCACCAGCTCTCGGACGAACAGCTCTCGG  123509
1176    ...............LysArgLeuLeuAspAlaGlnLnArgGlySe  1185
123508  CGGTAAATTACTTCAGTATGCAGCGGCTGCCAGCTCCTGGAGCGGGTATT  123459
        :::||||||      :::|||
1185    rPheProProTrpValGlnThrArgValArgTrpProHisLeuGlyT  1202
123458  GCCACCAAATTGGAGGCTCCAACGCTCGC.........TCAGCGCCGGCAG  123418
        ::::
1202    hrAlaTrpProProGln...........  1207
123417  TAGCTCCTCCACCTCAGCATCGTGGTCCACCTCACCTTCGCAAGAGCG  123368
        :::||   |||||||||
1208    ...ProHisProProGlyCysArgSerGlnArgThrAlaSerSe  1223
123367  GCGGCTCATCCGCCTACGGGCGGCTGAATGGCCGCTGAAGCGCGAGGAT  123318
        ||||||||||      :::|||
1223    rGlyThrProGlnThrThrSerProPro...........SerIleThrA  1236
123317  GTCTACGGCAGTGGCTACGGCGGCAGCAGTGGAAAATGTGGGCTATGGCTT  123268
        :::||||       :::|||:::
1236    spLeuLeuPro.Cys......AlaProTyrProGlyProProSerValTh  1250
123267  GTCCACTTCCAGTGCCGCGGGACACCATCAACATCCCACCAGCAGCAGC  123218
        ||||||||||     ||||:::||
1250    rAspLeuLeuProSerAlaProAspProGlyThrSerProThrThrArgT  1267
123217  AGAATCCGATGCAGCAGCAGCAAAGGGAACGGGATCAGGAACACAAG  123168
        :::|||     ||||||||
1267    hr..........GlyArgHisGlnGlyThrAlaGlyLeuSerPro  1278
123167  CAGTATGCCGGCAGTGTGCGGGCAGCGTGGGATCGGCCACATCAGCGGC  123118
        |||||||      |||:::
1279    LeuThrProPro.........TrpGlyThrHis  1286
123117  TCAGAGGGCCCTGAGCTTGGACTTGGACTGGCGGCGCGCAC  123085
        ||     |||:::|||
```

Hernandez,J.R., Houck,J., Hostin,D., Houston,T.J.,
Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,L.,
Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.

```
TITLE       The genome sequence of Drosophila melanogaster
JOURNAL     Science 287 (5461), 2185-2195 (2000)
MEDLINE     20196006
REFERENCE   2  (bases 1 to 300994)
AUTHORS     Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
TITLE       Direct Submission
JOURNAL     Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
FEATURES             Location/Qualifiers
     source          1..300994
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="X"
     mRNA            complement(join(<16198..16317,16386..16574,16635..16845,
                     16915..17053,17129..17411,17526..17607,17676..17975,
                     18107..>18483))
                     /gene="CG8527"
                     /product="CT24901"
                     /db_xref="FLYBASE:FBan0008527"
                     /db_xref="FLYBASE:FBgn0030844"
     gene            complement(16198..>18483)
                     /gene="CG8527"
                     /map="16B4-16B6"
                     /db_xref="FLYBASE:FBan0008527"
                     /db_xref="FLYBASE:FBgn0030844"
     CDS             complement(join(16198..16317,16386..16574,16635..16845,
                     16915..17053,17129..17411,17526..17607,17676..17975,
                     18107..18483))
                     /gene="CG8527"
                     /note="CG8527 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008527"
                     /db_xref="FLYBASE:FBgn0030844"
                     /protein_id="AAF48734.1"
                     /db_xref="GI:7293356"
                     /translation="MPQEKRHPPESRTQRFLETLVIFRRSLIYQTKEFFQNSTLHGVR
                     YIAESGRPIGEKFMWFCFTSIGAVTALVIIMSLWEKFQTNPTITGLDTDFHNQNVVFP
                     TTVVCPEAAFDHDKTYERKVYNTLANYDEAQAQMYTPFLRILTLSLNFENVRDAKVLSQS
                     IPQNLLDAHTIREWAFEGHIDCKNVFVSCKYRDEDIPCCDHFEPIYTEHGCYAFNSR
                     FKSTPTEDVKTGAPHDLYETDKKWALFFIPNSTSRIFIFSNEEYFGSDFNAQIDWSEP
                     QLVEVRISKKNTYTTDDARQLSIGQRKCIFSDEVKLNYFPDAYTFSSCMKQCRMNKAI
                     KLCKNPPFYKPIPNVPMCSIKDFDCLDEFKSNITNIKDCLQCELSCSKTVFNIDKLI
                     KMSDRPESLGVLVEFLTWPIIRYKREVLFGWDLLVSFGGIASLFGFSLLSGVEIIY
                     YFTLRACCMVYKNRQELYEIEEKIRQPPKIDLKLSLKSHNPRIFGDTPTSAVLKVKP
                     AEESAGPNPLGGYQRKHEKDYSLNTKNYKTPKVVPDYGYTSKTKNSWTTYDHSQYMP
                     "
     mRNA            complement(join(21273..21792,21868..21999,22058..22191,
                     22257..22823,22889..23887,24168..24320,24398..25116,
                     25191..25449))
                     /gene="CG8465"
                     /product="CT24743"
                     /db_xref="FLYBASE:FBan0008465"
                     /db_xref="FLYBASE:FBgn0030845"
     mRNA            complement(join(21273..21792,21868..21999,22058..22191,
```

```
seq_name: gb_in1:AE003506

seq_documentation_block:
LOCUS       AE003506   300994 bp   DNA   INV   04-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386053 section 23
            of 30, complete sequence.
ACCESSION   AE003506 AE002593
VERSION     AE003506.1  GI:7293355
KEYWORDS    HTG.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 300994)
AUTHORS     Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
```

```
              22257. .22823,22889.      .23887,24168. .24320,24398. .25116,
              25191. .25370,27861.     .28246))
              /gene="CG8465"
mRNA          /product="CT24757"
              /db_xref="FLYBASE:FBan0008465"
              /db_xref="FLYBASE:FBgn0030845"
              22257. .22823,22889.     .23887,24168. .24320,24398. .25116,
              25191. .25370,27861.     .28115,28453. .28555))
              /gene="CG8465"
              /product="CT24733"
gene          /db_xref="FLYBASE:FBan0008465"
              /db_xref="FLYBASE:FBgn0030845"
              complement(<21273. .>28555))
              /gene="CG8465"
              /map="16B4-16B7"
              /db_xref="FLYBASE:FBan0008465"
              /db_xref="FLYBASE:FBgn0030845"
CDS           complement(join(21399. .21792,21868. .21999,22058. .22191,
              22257. .22823,22889.      .23887,24168. .24320,24398. .25116,
              25191. .25370,27861.     .28092))
              /gene="CG8465"
              /note="CG8465 gene product [alt 1]"
              /codon_start=1
              /db_xref="FLYBASE:FBan0008465"
              /db_xref="FLYBASE:FBgn0030845"
              /protein_id="AAF48735.1"
              /db_xref="GI:7293357"
              /translation="MSTYFGVYIPTSKAGCFEGSVSQCIGSIAAVNIKPSNPASGSAS
              VASGSPSGSAASVQTGNADDGSAATKYEDPDYPPDSPLWLIFTEKSKALDILRHYKEA
              RLREFPNLEQAESYVQFGFESIEALKRFCKAKPESKPIPIISGSGYKSSPTSTDNSCS
              SSPTGNGSGFIIPLGSNSSMNLLLSDSPTTSSPSSSSNVIANGRQQOMQQQQQPQQ
              PDVSGEGPPFRAPTKQELVEFRKQIEGGHIDRVKRIIWENPRFLISSGDTPTSLKEGC
              RYNAMHICAQVNKARIAQLLLKTISDREFTQLYVGKKGSGKMCAALNISLLDYYLNMP
              DKGRGETPLHFAAKNGHVAMVEVLVSYPECKSLRNHEGKEPKEICLRNANATHVTIK
              KLELLLYDPHFVPVLRSQSNTLPPKVGQPFSKDPPNLQHKADDYEGLSVDLAISALA
              GPMSREKAMNFYRRWKTPPRVSNNVMSPLAGSPFSSPVKVTPSKSIFDRSAGNSSPVH
              SGRRVLFSPLAEATSSPKPTKNVPNGTNECEHNNNNVKPVYPLEFPATPIRKMKPDLF
              MAYRNNNSFDSPSLADDSQILDMSLSRSLNASLNDSFRERHIKNTDIEKGLEVVGRQL
              ARQELEWREYWDFLDSFLDIGTTEGLARLEAYFLEKTEQQADKSETVWNFAHLHQYF
              DSMAGEQQQQLRKDKNEAAGATSPSAGVMTPYTCVEKSLQVFAKRITKTLINKIGNMV
              SINDTLLCELKRLKSLIVSFKDDARFISVDFSKVHSRIAHLVASYTHSQEVSVAMRL
              QLIQMLRSLRQLLADERGREQHLGCVCASLLLMLEQAPTSAVHLPDTLKTEELCCAAW
              ETEQCCACLWDANLSRKTSRRKRTKSLRAAAVVQSGQLQDTSGSTGSSALHASLGVG
              STSLGASRVVASASKDAWRRQQSDDEDYDSDEQVIFDCTNVTLPYGSSSEDEENFRT
              PPQSLSPGISMDLEPRYELFIFGNEPTKRDLDVLNALSNVDIDKETLPHVYAWKTAME
              SYSCAEMNLNVKVQKPEPWYSGTSSSHNSQPLLHPKRLLATPKLNAVVSGRRGSGPLT
              APVTPRLARTPSAASIQVASETNGESVGTAVTPASPILSFAALTAATQSFQTPLNKVR
              GLFSQYRDQRSYNEGDTPLGNRN"
CDS           complement(join(21399. .21792,21868. .21999,22058. .22191,
              22257. .22823,22889.      .23887,24168. .24320,24398. .25116,
              25191. .25370,27861.     .28092))
              /gene="CG8465"
              /note="CG8465 gene product [alt 2]"
              /codon_start=1
              /db_xref="FLYBASE:FBan0008465"
              /db_xref="FLYBASE:FBgn0030845"
              /protein_id="AAF48736.1"
              /db_xref="GI:7293358"
              /translation="MSTYFGVYIPTSKAGCFEGSVSQCIGSIAAVNIKPSNPASGSAS
              VASGSPSGSAASVQTGNADDGSAATKYEDPDYPPDSPLWLIFTEKSKALDILRHYKEA
              RLREFPNLEQAESYVQFGFESIEALKRFCKAKPESKPIPIISGSGYKSSPTSTDNSCS
              SSPTGNGSGFIIPLGSNSMSMNLLLSDSPTTSSPSSSSNVIANGRQQOMQQQQQPQQ
              PDVSGEGPPFRAPTKQELVEFRKQIEGGHIDRVKRIIWENPRFLISSGDTPTSLKEGC
              RYNAMHICAQVNKARIAQLLLKTISDREFTQLYVGKKGSGKMCAALNISLLDYYLNMP
              DKGRGETPLHFAAKNGHVAMVEVLVSYPECKSLRNHEGKEPKEICLRNANATHVTIK
              KLELLLYDPHFVPVLRSQSNTLPPKVGQPFSKDPPNLQHKADDYEGLSVDLAISALA
              GPMSREKAMNFYRRWKTPPRVSNNVMSPLAGSPFSSPVKVTPSKSIFDRSAGNSSPVH
              SGRRVLFSPLAEATSSPKPTKNVPNGTNECEHNNNNVKPVYPLEFPATPIRKMKPDLF
              MAYRNNNSFDSPSLADDSQILDMSLSRSLNASLNDSFRERHIKNTDIEKGLEVVGRQL
              ARQELEWREYWDFLDSFLDIGTTEGLARLEAYFLEKTEQQADKSETVWNFAHLHQYF
              DSMAGEQQQQLRKDKNEAAGATSPSAGVMTPYTCVEKSLQVFAKRITKTLINKIGNMV
              SINDTLLCELKRLKSLIVSFKDDARFISVDFSKVHSRIAHLVASYTHSQEVSVAMRL
```

```
              QLIQMLRSLRQLLADERGREQHLGCVCASLLLMLEQAPTSAVHLPDTLKTEELCCAAW
              ETEQCCACLWDANLSRKTSRRKRTKSLRAAAVVQSGQLQDTSGSTGSSALHASLGVG
              STSLGASRVVASASKDAWRRQQSDDEDYDSDEQVIFDCTNVTLPYGSSSEDEENFRT
              PPQSLSPGISMDLEPRYELFIFGNEPTKRDLDVLNALSNVDIDKETLPHVYAWKTAME
              SYSCAEMNLNVKVQKPEPWYSGTSSSHNSQPLLHPKRLLATPKLNAVVSGRRGSGPLT
              APVTPRLARTPSAASIQVASETNGESVGTAVTPASPILSFAALTAATQSFQTPLNKVR
              GLFSQYRDQRSYNEGDTPLGNRN"
CDS           complement(join(21399. .21792,21868. .21999,22058. .22191,
              22257. .22823,22889. .23887,24168. .24320,24398. .25116,
              25191. .25410))
              /gene="CG8465"
              /note="CG8465 gene product [alt 3]"
              /codon_start=1
              /db_xref="FLYBASE:FBan0008465"
              /db_xref="FLYBASE:FBgn0030845"
              /protein_id="AAF48737.1"
              /db_xref="GI:7293359"
              /translation="MPTHQHCHRHDGAADSPLWLIFTEKSKALDILRHYKEARLREFP
              NLEQAESYVQFGFESIEALKRFCKAKPESKPIPIISGSGYKSSPTSTDNSCSSSPTGN
              GSGFIIPLGSNSMSNLLLSDSPTSSPSSSSNVIANGRQQOMQQQQOPQQPDVSGE
              GPPFRAPTKQELVEFRKQIEGGHIDRVKRIIWENPRFLISSGDTPTSLKEGCRYNAMH
              ICAQVNKARIAQLLLKTISDREFTQLYVGKKGSGKMCAALNISLLDYYLNMPDKGRGE
              TPLHFAAKNGHVAMVEVLVSYPECKSLRNHEGKEPKEICLRNANATHVTIKKLELLL
```

```
alignment_scores:
       Quality:  1257.50        Length:     1432
       Ratio:       1.642       Gaps:         57
Percent Similarity: 53.492   Percent Identity: 29.190

alignment_block:
US-09-294-298-6 x AE003506

Align seg 1/1  to: AE003506    from: 1   to: 300994

   19 PheLeuSerArgArgLeuLysSerSerIleLysArgThrLysSerGlnPr  35
      :::|||   ||:|||   ::|||||||||||||||||||||||
195821 TTCCTTGCCACAGA...TCCATCCTTGAAGCGGACCAAGTCCGTGAC 195867

   35 oLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheArgS  52
      :||||:::||:||||::
195868 CAAGCTGGAGCGGACCAAG...                             195886

   52 erAlaAspHisAspArgArgLeuMetGlnSerPheLysGluSerHis  68
      |||        :::::||:    ||::::||
195887 ...CGCGGATCCGGCGGACTGAGGGGCTCCGC 195916

   69 SerHisGluSerLeuSerProSerSerAlaAlaGluAla........  82
      ||||||||||||::|||
195917 TCGCACGAGAGTCTGCTGTCCAGTCACGCCGTCCACCATAGGTGA 195966

   82 ...........................................  82
195967 GTTCTGGCTCATAATGAGTTTTGGCCAAACCTCAAACCATTTCTTT 196016

   83 ...........LeuGluLeuAsnLeuAspGluAspSerIleIleLysPro  95
      ||||||:::
196017 TCTTTCTTCTTCTAGATCTCTCCGCACTGGGCGCGGTGGGCGTGGCGCCC 196066

   96 ValHisSerSerIleLeuGlyGlnGluPheCysPheGluValThrThrSe 112
      ||||||:||::|||:||||||   ||||||||::
196067 GTGCATCACGGTTCTGGGACCTCGTCCACTGTTTCCAGGTACGGGCGCGG 196116

  112 r...SerGlyThrLysCysPheAlaCysSerAlaAlaGluArgAspL 128
      ||       ::::||| |||  ||||||||||
196117 GCCTCGTGCGCAGGCTACTACTCATGCGGATCCGCCGCCAGGAGCGCGACC 196166

  128 ysTrpIleGluAsnLeuGlnArgAlaVallysProAsnLysAspAsnSer 144
      ||||     :::::||||     |||:|||
196167 TTTGGATCTACTCGCTCCGCAAGTCGATCGCTCCGAATGCAGAGCACACG 196216

  145 ArgArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGluE 161
      ||||||:::||||||   ||||||::||:::||||   ||||||:|||
196217 CGTCGCACGGACAACTCCGCTGAAGATGTGGGTACGAGGCCAAAAATCT 196266
```

us-09-294-298-6.rge

```
161    uProProLysArgTyrTyrCysGluLeuCysLeuAspMetLeuT          178
       ||||||||||||:||:||||||:|||||||    ||||
196267 GCCGCCCAAGAAGCGTTACTTTTGCAACTGCAATTGGACAAGACGCTGT      196316

178    yrAlaArgThrThrSerLysProArgSerAlaSerGlyAspThrValPhe    194
       ||::|||||   |||   :::|:||  |||   :::|||
196317 ACGGCCGACTTCGGTGAAGCTGCAGACG.......GATCTGCTGTTT       196357

195    TrpGlyGluHisPheGluAsnAsnLeuProAlaValArgAlaLeuAr      211
       :::|||||||||:||||:    ||||||||||:|||||:|  |||
196358 TGGGGGGAGCACTTCGATTTCCCGACATACCCGAGATTAATGTGATCAC    196407

211    gLeuHisLeuTyrArgAspSerLysLysArgLysLysAspLysAlaG     228
       |::|||   :|||:||:   :||||||||||||:||||:||
196408 TGTTAACGTTTTCCGTGAGGTGGACAAGAAGAAGAAGCGGGACAAATACC   196457

228    lyTyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgHis   244
       :::||||   |||   ||||   :|::||   ::|:|||
196458 AATTTGTGGGATCCGGTGAAGATACCCGTGACGATGTCACCTCCAGATTG   196507

245    PheThrGlyGlnTrpTyrProValThrLeuProThrGlySerGlyValGlySe  261
       ||||||||||| |||    :::||||
196508 CCCTGCGAACAATGGTATCCCATA...CTGAGCGACAAGGCGGGCGACAG  196554

261    rGlyGlyMetGlySerGlyGlyGlyGlyGlyGlyGlyGlySerGlyG      278
       |||||||||||
196555 TCTGGGCAGGACATCGGGC.............GCGGCCGCAGTGGGT      196589

278    lyLysGlyLysGlyLyCysProAlaValArgLeuLysProAlaAlaArgTyrGln 294
       :||| |||| ||::||||::||:|  ||   :||
196590 CCAAGGACAAGGACAATTGCCCACGTGAGATCAAGTGTCGTTTCCAG      196639

295    ThrMetSerIleLeuProMetGluLeuTyrIleLysGluHisAlaGluTyrVa  311
       :::||||||||  |||   :|||||   |||
196640 AGCACCGACATCCTGCCCATCAATGTTACGGCAACTTTTTGACGTACCT    196689

311    lThrAsnHisTyrArgMetCysLeuAlaLeuAlaLeuGluProAlaLeuAsnV  328
       ||::||| ::|   |||   ::||||||   :|:|||||
196690 CAAGGAGAACTATAAGCGGTGTGCAGAACCCTGGAGACCCGGTGATCGGAG  196739

328    aLLysGlyLysGluGluValAlaSerAlaLeuValHisIleLeuGlnSer   344
       ||||::||||:|||  :::|||||   ::|:|||||
196740 TCAAGGCCAAGGAAGAGGAGCCATTGGACAGGCACTGGTGCTGATGCACGCA 196789

345    ThrGlyLysAlaLysAspPheLeuSerAlaMetSerGluValAs        361
       ||   |||   :||||||:   :||||||
196790 CAGGGATTGGCGGGCGCCTTCCTCACCGATGGTGCCCCTCGATCTGCT     196839

361    pArgPheMetGluArgGluHisLeuIlePheArgGluGluAsnThrLeuAlaT 378
       |||:    ::||||:   |:|| |||   :||||||||
196840 GCGAGTTGGCGAT...CAGAGGCTTACGTTCAGGGCAACTCCTTGGCCA    196886

378    hrLysAlaIleGluGluLeuMetArgLeuIleGlyGlyLysTyrLeuLys   394
       ::|||   :|||||   ||||:   |||||||||
196887 CCAAGAGCATGGAGGAGCCATTCCTCAAGCTGACGGCCGAACAGTATCGCAG 196936

394    AspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsnCy   411
       ||:|||   ||||   |||   ::||||||::|||:
196937 GACACACTATCCGCACCAATAAACGAGCTAATTCAGTCGGAGAGGGACTG   196986

411    sGluValAspProIleLys.........CysThrAlaSerSerLeuAlaG   425
       ||||||||| |||    :::|||||
196987 CGAGGTGGATCCCACCAGACGGTTCGGCGGGTTCGCTGCAGC           197036

425    luHisGlnAlaAsnLeuArgMetCysCysLeuLeuAlaLeuCysLysVal   441
       :::||||||||| ||||    :::|:||||
197037 GACAGCAGGCCGCCTTGCGTGCCGCGGTCCGAGGGGCGGTGCAGTGCATC   197086

442    ValAsnSerHisCysValPheProArgArgLeuLeuLysGluValPheAlaSe 458
       ::::||||   :|||   :|||||   ||||||
197087 TTCGAATCGCACAAGCACATTTCCCGCCCAGTTACCGAAAATTGCTTTGCGAC 197136
```

```
458    rTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAspArgL    475
       ::: :||||||||| ||   |:::|||||:::||
197137 GTTCCGGCGGAGCGCTTGCAGCAGCTGGCCGTCAGGATAGGCGACAAACC     197186

475    euIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSer    491
       ||||||||||||:|||||:|||:::|:|||::|||
197187 TGATCTCGGCGGAGCCATTTTCCTGCGCTTCCTGTGCCCGGCCATCCTGTCG  197236

492    ProSerLeuPheGlyLeuMetGlnGlyTyrProAspGluGlnThrSerAr    508
       ||||||||||:   ::||||:|||   :::||:||::||||
197237 CCGTCGCTCTTCAATATCACCAGGGAACTGCCGTCCGACGGGCCACCCG      197286

508    gThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerL    525
       |::||||||:||||:||||||||||||||||||::|||||
197287 CAATCTCACACTGGTGGCCAAGACCCTGCAAACATTGGCCAATTTCACCC    197336

525    ysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGlu    541
       ::||| :::||||   :||||||||||||::::||||:::|||::|
197337 GCTTCCAGGGCAAAGAAGACTTTATGGAGTTTCCAACGATTTCCTCGAG    197386

542    LeuGluTrpGlySerMetGlnGlnLeuTyrGluIleSerAsnLeu..      557
       :::||||||||||||||||:||||   ::
197387 CAGGAGGCCGCTGCCATGCAACAGTTTCTGGAGATTATATCCACGCGGCC   197436

558    .........AspThrLeuThrAsnSerSerPheGluGlyT             569
                |||    ::: |:||||
197437 GGAGCACCCAGCTCCAGACTCCGATCCTCGAT.......TGGGCCGGGT    197477

569    yrIleAspLeuGlyArgGluSerThrLeuHisAlaLeuLeuTrpGlu      585
       ||||   ::|   :::|||||:::|||:||   ::|||
197478 ACATCGACCAGGGCAAACAGTTGTCCATACTACACAGTTGTCTCACCGAA   197527

586    ValLeuProGlnLeuSerLysGluAlaLeuLeuLeuLysLeuGlyProLeuPr 602
       ||| |:|||:   |:||||:||||||||
197528 AGCCTGGCCAAGCTGCCGGAGGCCAGCAGCGAGCTGGATCCGTTGCA      197577

602    oArgLeuLeuSerAspIleSer........ThrA                   611
       ::::||:::||:|||||||
197578 GCATATTCTCGATGAAATCAGCCGACCCAAAGAGCATGGCATGGGCACAG   197627

611    laLeu..........                                      612
       |||||
197628 CACTGCGCGGTGGATATTTGCGGCGCCACCTCGTCTACGCACTCGATAGCC  197677

613    .........ArgAsnPro..........                         615
                ||||||||
197678 AGCGAGAATCAGGAGAATCCAATCCGGGATCCTCCGGCGTCCCACGCTGG   197727

616    ....AsnIleGlnArgGlnProSerArgGlnSerArgGlnSerArgSerG   631
           || |||||||||
197728 CTCCAACTCGGAGCAGTTACTGCCACAACAAAGCCAGTTGGCCCAGCCGC   197777

631    lnProMetValLeuArgGlyPro...SerAlaGlu........MetGln   643
       || ||||||||  ||||||||
197778 AGCAGTGGCGATTGTTAGTAAACCATTATCTCGCGGAGCGCGGCCATCATGCGGA 197827

644    Gly.TyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerPheM  660
       ||||||  ||::||||||||||||
197828 GGAGTACTTACGCCGAATTCTCTGGAGAAGAATATCTTTAGATACAATGA   197877

660    etAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerPro  676
       :: :::
197878 TCCCACCGTTAATG.............................          197891

677    ThrLysGluLysProProProProProGlyGlyLysAspLeuPh        693
       ||||
197891 ..............................................       197891

693    eTyrValSerArgProLeuAlaArgSerProAlaTyrCysThrS        710
       ||||
197892 ..........................GCTTACTGCAGCA              197904

710    erSerSerAspIleThrGluProGluGlnLysMetLeuSerValAsnLys  726
```

198517 CAAACTGAATTGGGAATACCCGATCACTCAGGTGGCTATGTGCGGGGTA 198566
         ||| |||||||||||||||||||||||||||||||
993   ......GlnGlnSerLeuSerLysGluGlySerIleGlyGlySerGlyG 1007

198567 ATAATTTGAATCCCAACTCGAATATGCCCAAGAACTTGGAGGATCTGGAC 198616
                              ||||||||||||||||||
1007 lySer...........GlyGlyGlyGlyGlyGlyLeuLysPro 1018

198617 GATCTGTTCAAGTACGCCGAGGAGCATGAGCGTGGCGGGACCAGCGAACCA 198666
        |||||| ::|||
1019 SerIleThrLysGlnHisSer...GlnThrProSerThr.LeuAsnPro. 1033

198667 TCACAATCACAACCAGGTCAGCAGAACCACCAGGGACATCTGAAGCCGCCG 198716
        :::||||||||||::|||:::|| |||::||::||
1034 ..ThrMetProAlaSerGluArgThrValAlaTrpValSerAsnMetPro 1049

198717 CCGCCGGTTCCCGGCCAAGGAGCAGCGTGTCGGCGAAAAGCAGT..... 198757
         ::::::||:::||||:::||| ||| |||
1050 HisLeuSerAlaAspIleGluSerAlaHisIleArgGluGluTyrArLy 1066

198758 CACTGCAGTTCTGGCTACCAGAGCATCTCCACAAATCCC.......... 198796
         ::||||||
1066 sLeuLysGluTyrSerLysSerMetGluSerArgLeuAspArgValL 1083

198797 ...........TCGCCCTCGCAGTCTCCAGTCCCGTGGAG......... 198826
                   |||::|||:::::||::::||
1083 ysGluTyrGluGluGluIleHisSerLeuLysGluArgLeuHisMetSer 1099

198827 ..........AGCCAGCTGAAGGCCCGATGGGCAGTCAC.......... 198856
          ::::||||
1100 AsnArgLysLeuGluGluTyrGluArgLeuLeuSerGlnIleGluGluGl 1116

198857 AATGCGCCGCTG...GCCTTCAAGAATCCCTCCTATCAGCTTCAGCCCCA 198903
        ||| ||| ||| ::: ||||
1116 nThrSerLysIleLeuMetGlnTyrGlnAlaArgLeuGlnLnSerGluL 1133

198904 AACTGGCCTCGTCCAGATCATCGGCACGAGTAATACACCAGCAGCAGC 198953
        |||::|||::::||| |::: :: ||||||::||
1133 ysArgLeuArgGlnGlnGlnValGlnLysAspSerGlnIleLysSerIle 1149

198954 AGCAGCAGCAACAACAACAGTTT......GGCAGCCGCTTGAAACCAATT 198997
        :::|||||:| :::|||::::::||
1150 IleGlyArgLeuMetLeuvalGluGluLeuArgArgAspHisProAl 1166

198998 GGCGGCCGGACTGCTGGCCGGAGGGCGGCGTTTCCTCAACAGTGGCGAGC 199047
        ||| |||::: ||:::||||||
1166 aMet.AlaGluProLeu........ProGluProLys........ 1175

199048 CTTGGAGGGCGGCCACTTTGACGCCCAGCTCCTCGACGAACAGCTCTCGG 199097
        ::: ||::: ::: |||:::
1176 ....LysArgLeuLeuAspAlaGlnArgGlyse 1185

199098 CGGATAATTACTTCAGTTATGCAGCGGCTGCGAGCTGCTGGAGCGGGTATT 199147
        |||||||| |||::||
1185 rPheProProValGlnGlnThrArgValArgTrpProHisLeuGlyT 1202

199148 GCCACCAAATTGGAGGCTCAACGCTCGC........TCAGCGGCGGCAG 199188
        ::::||| ||||
1202 hrAlaTrpProProGln............................. 1207

199189 TAGTCTCCCACCTCAGCACTCTGCTCCACCTCGAATCTTGGCAAGAGCG 199238
        :::|||||| ::||||||
1208 ...ProHisProProHisProGlyCysArgSerGlnArgThrAlaSerSe 1223

199239 GCGGCTCATCCGCCTACGGCGCGCTGCGTGAATGGGCCGCTGAAGCGCGAGGAT 199288
        |||||||| :::|||
1223 rGlyThrProGlnThrThrSerProPro.............SerIleThrA 1236

199289 GTCTACGGCAGTGCTACGGCGCAGCAGTGGAAATGTGGGCTATGGCTT 199338
        ::::||| :::|||
1236 spLeuLeuPro.Cys........AlaProTyrProGlyProProValTh 1250

199339 GTCCACTTCCAGTGCCGCCGGACCACCATCCCACCACCAGCAGC 199388
        |||||||| :::|||

197905 ||| |||:::: AGCAGCAACAAC 197936
        ||||||||::   :::::::
727   SerValSerMetLeuAspLeuGlnLeuGlyAspGlyProGlyArgLeuAs 743

197937 AGCAACAGCAACATC.............AGCAGCTGCAA 197962
        ||| ||| :::
743   nSerSerValSerAsn.........LeuAlaAlaValGly..AspLeuLe 757

197963 CAGCATGGCCATCAGCAACAGCCGCACCACCAGCATCCACTCCAGATGCT 198012
                                           :::|||
757   uHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaP 774

198013 CTCCAATTCACAAACCTCCATT..GCCGGCAACCAATATATGAGTTCGC 198059
        | :::||||||::|||:::   :::|
774   roAlaGlyArgLeuSerGlnIleGlySerGlySerIleThrAlaAlaGly 790

198060 CAGGAGGCCTGCAGCATGCCCAATCGCAGACCTCGATGGCGTCCTCATCG 198109
        ||| |||:::|| |||
791   MetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnGl 807

198110 CTT.. 198112
        :::
807   nLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaA 824

198112 198112
824   spGlyProGlyProProAlaGlyHisGlyGlySerGly.....His 838

198113 ..........AATGGGAGCAGCAGCAATTTGCTGCAC 198139
          :::|||::::|||:::||
839   GlyProProSerHisHisHisHisHis...........Hi 851

198140 GGC..CACCAGCAGCATGCCCATCACCCGGCAGCAACTGCA 198177
        ||| |||||
851   sHisHisArgGlyGlyGluProGlyAspThrPheAlaProPheHisG 868

198178 TCCACATCACTGC......CCGCCG............ 198196
868   lyTyrSerLysSerGluAspLeuSerThrGlyValProLysProProAla 884

198197 ...............GCGCCACAGACCAGTGCC 198214
                        |||
884   ....ThrValSerAlaAlaGlnLysProAr 896

198215 TCCAGCACTATGGAGGCGATCGCATGGAACTATCCGTATATGTCCCA 198264
        |||:::||| |||:::||
896   AlaSerIleLeuHis...............SerHisSerTyrSerAspGl 913

198265 C......AATGGCAATGACTACGAG 198283
        :
913   spAsnLeuGlnHisMetLeuSerProProGlnIleThrIleGlyProGln 929

198284 ...........ACCACACGCCTTCGAGCACTCGCTCCAGGACA 198316
                    ||| ||::||
929   ....GlyGlyGlyGlySerGlyGlyGlyGlySe 944

198317 CTGCCACGGAATGGAAATCCCAATGCCAATGGCAACGTGGGCAGCAGCAA 198366
        :::|||||| |:|:|||||||:::|||
944   rGlyGlyGlyGlyGlnArgGlyGlnLeuGlnLeuSerLeuGlnA 960

198367 CAATAACCAGAGCGGCCAGCTACGATGGACATGCACGCAGGTTCCAAATCC 198416
        ::||||||
960   ....ThrSerAlaAlaGlnLysProGln 972

198417 AGATCTCTGGGTTCGATACGACAGCAGTGCTTTTGTCTGCAAGTCGCCCACA 198466
        :::|||:::||
972   GlnGlnLeu.................... 982

198467 CCCATGATGAAATCCCAGTTTGGGACCAGCGGCGGGGCCGGCGAAGTCATCA 198516
        ||| ||||||:::
982   gPro................Ser.SerGlyAsnLeuLeuGlnSerPro 992

198516 GTCCACTTCCAGTGCCGCCGGACCACCATCCCACCACCAGCAGC 198516
        :::   :::|||
983   GluProSerTyrGlyProAlaArgProArg......... 992

```
 1250  rAspLeuLeuProSerAlaProAspProGlyThrSerProThrThrArgT  1267
       :::|||             |||||||||  ||:::|||
199389  AGAATCCGATGCAGCAGCAGCACCAAAGGAACGGATCAGGAACACAAG  199438

 1267  hr..........GlyArgHisGlnGlyThrAlaGlyLeuSerPro  1278
       ::          |||  :::::::|||  |||  ::::|||
199439  CAGTATGCCCGCAGTGTGGCGGCAGCGTGGGATCGGCCACATCAGCGGC  199488

 1279  LeuThrProPro........TrpGlyThrHis  1286
       |||             |||  |||::::|||
199489  TCAGAGGGCGCCTGAGCTTGGACTCGGCGCGCAC  199521
```

seg_name: gb_in1:AB011280

seq_documentation_block:
```
LOCUS       AB011280     3984 bp    mRNA            INV       25-FEB-1998
DEFINITION  Caenorhabditis elegans mRNA for GAP-2-4, complete cds.
ACCESSION   AB011280
VERSION     AB011280.1   GI:2911175
KEYWORDS    GAP-2-4.
SOURCE      Caenorhabditis elegans (strain:Bristol N2) mixed stage
            hermaphrodite cDNA to mRNA.
ORGANISM    Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
            Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 3984)
AUTHORS     Yamamoto,M., Iino,Y. and Hayashizaki,S.
TITLE       Direct Submission
JOURNAL     Submitted (18-FEB-1998) to the DDBJ/EMBL/GenBank databases.
            Masayuki Yamamoto, University of Tokyo, Department of Biophysics
            and Biochemistry; 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
            (E-mail:myamamot@ims.u-tokyo.ac.jp, Tel:81-3-3812-2111,
            Fax:81-3-5802-2042)
REFERENCE   2  (sites)
AUTHORS     Hayashizaki,S., Iino,Y. and Yamamoto,M.
TITLE       Characterization of the C. elegans gap-2 gene encoding a novel
            Ras-GTPase activating protein and its possible role in larval
            development
JOURNAL     Genes Cells (1998) In press
FEATURES    Location/Qualifiers
source      1..3984
            /organism="Caenorhabditis elegans"
            /strain="Bristol N2"
            /db_xref="taxon:6239"
            /dev_stage="mixed stage"
            /sex="hermaphrodite"
gene        20..3382
            /gene="gap-2"
CDS         20..3382
            /gene="gap-2"
            /codon_start=1
            /product="GAP-2-4"
            /protein_id="BAA24960.1"
            /db_xref="GI:2911176"
            /translation="MHPSPIPQEYQPQANYSSRIHVHFAQSQSRRPFQRDQRNENASI
            RKRLSRSCDHLSDHNFSPRLPPIRRKPTKHVPSTLSLVIHDAIHDETNSGGAQVGGES
            LDRRGWSGASTGNRRTLTASVHNNLMNGRMSSSSHNLSTRLSGSTQNLNQPTNAYGFP
            IVPPANNPSDSSASGRIANLLSRPFRSNPLKRTKSVSKMEKSLAEANQHSLHRVDASN
            TPSRDSSLYAQPPARRHLSQPAREGSLRACRSHESLLSSAHSTHMIELNEDNRLHPVH
            PSIFEVPNCFRLASTYYSCRTPLERAKWMENLRKTMNPRRDQQRRTENSMLIWILEAK
            GLPAKRKYYCEMTLDKTLYAKTSSKARTDNVFWGENFEFMMLPKIDEVCVSLFRESDS
            KKKDTLIGYVTGIDLQLSSRSPVERWYTVNTSHSDSGTSRIASALGGKSSSQESPSL
            RIKARWOSVHILPLRAYDNLLETLCYNYLPLCEQLEPVLNVRDKEDLATSLVRVMYKH
            NLAKEPLCDLIMKEVEKLDNDHLMPFRGNTLATKAMESFMKLVADDYLDSTLSDFIKTV
            LQCEDSCEVDPQKLGNVSNSSLEKNRALLMRYVEVAWTKILNNVHQLPKNLRDVFSAL
            RCRLEAQNREALADTLISSSIFLRFLCPAILSPSLFNLVSEYPSPTNARNLTLIAKTL
            QNLANFSKFGCGKEPHMEFMNEFVDREWHRMKDFLRISSESKSGPEKNADAIVDAGKE
            LSLIATYLEEAWTPLLQEKNGNKHPLSNVKSVLSELAECKRRSDNGVFHSPMVQQPSS
            DYENSPQOHVVPRHENVPAYRSTPPTGQATVLGRSTNRPATHLLTSDDYVLSSAFQTP
            SLRPGGTRLSDETGTSSRTSDKTTSSAEIRDDTDSDFELREDRGRGGNRRKRLPRTD
            ASPSSSOQASSGYLSNNPSRSYSNSSSSPVERMAALSIANPCFGPGPSSGYAIPAE
            PKEIVYQKRASPPPYDPDVHNYHYQPMQYYAVPPDCQVSPRTQATGGVNAQNRLSLPR
```

```
                         TNPRASRNSTLLRPSVVNVPDDWDRTSDYWRDRGENNYRSQLESQVESQAREIERLMR
                         ENIELKSKRCLQQKLWIFPSDLTVLVRIPTIL"
polyA_site               3984
                         /note="15 A nucleotides"
BASE COUNT    1167 a    943 c    848 g   1025 t      1 others
ORIGIN
```

alignment_scores:
```
            Quality: 1173.50                  Length:    909
            Ratio:     2.122                    Gaps:      27
Percent Similarity:  60.836        Percent Identity:  33.883
```

alignment_block:
US-09-294-298-6 x AB011280

```
Align seg 1/1  to: AB011280  from: 1  to: 3984

 11  PheProPheAlaProGln.................................... 16
     |||||||:::::|||
494  TTCCCCATCGTGCCACCGGCAAATAACCCATCGGATTCGTCTCGCTGCGTCAGG  543

 17  ......GlnGlyPheLeuSerArgArgLeuLysSerSer...IleLysA  30
          :::|||||||||                 :::||||    ::::|||
544  TCGTATTGCAAATCTCTTATCACGACCGTTCCGTTCAAACCCACTAAAAA  593

 30  rgThrLysSerGlnProLysLeuAspArgThr..................... 40
     ||||||||||  |||:::||
594  GAACAAAGAGTGTGAGCAAAATGGAAAAATCGCTGGCAGAAGCAAATCAA  643

 41  ...SerSerPheArgGlnIleLeuProArgPheArgSerAlaAspHisAs  56
        |||           |||           |||  |||   ::
644  CACTCTTCTTCACAGAGTGGACCCCAGCAATACACCAGCAGAGTAGCAG  693

 56  p..............................ArgAlaArgLeuMetGln..  62
     :                               ::|||  |||
694  TCGTATGCACAACCTCCGGCTCGACGACATTTAAGCCAACCGGCTCGAG  743

 63  .....SerPheLysGluSerHisSerHisGluSerLeuLeuSerProSer  77
          |||:::::           ::|||||||||||||||||||||  :::
744  AAGGAAGTCTTAGGGCCTGCCGTTCACACGAATCATTGCTCAGTTCGGCA  793

 78  SerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAspSerIleIleLeLy  94
     |||||||||::::||||||           ::|||||||||:::      :::
794  CATTCAACACATATGATTGAATTGAAC......GAAGATAACCGCCTGCA  837

 94  sProValHisSerSerIleLeuGlyGlnGluPheCysPheGluValThrT  111
     ||||||||  |||||||||||||||||   ::||||||||||||||:::
838  TCCAGTACATCCTTCAATCTTTGAAGTTCCAAATTGTTTTCGACTGGCTA  887

111  hrSerSerGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAsp  127
     |||||                       :::||||||||||:::  ||||||
888  GCACC.............TACTATTCCTGTCGACACTGCCACTGAACGAGCA  925

128  LysTrpIleGluAsnLeuGlnArgAlaValLysProAsnLysAspAsnSe  144
     |||||||:::||||||||||           |||||||||||::: :::
926  AAGTGGATGGAGAACCTGCGCAAAACCATGAATCACGTAGGGATCAGCA  975

144  rArgArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGluL  161
     :|||||||::::|||  ::|||||||||||:::))|||:::
976  GAGACGTACTGAAAACAGCATGCTAATATGGATACTTGAAGCAAAGGGAC  1025

161  euProProLysArgTyrTyrCysGluLeuCysLeuAspAspMetLeu  177
     ||||           ||||||||||||||||||:::    ||||||
1026 TACCAGCCAAGAGGAAGTACTATTGCGAAATGACACTGGACAAGACGCTG  1075

178  TyrAlaArgThrSerLysProArgSerAlaSerGlyAspThrValPh  194
     ||||||||||:::||||||   |||:::      ||::::::||||||
1076 TATGCCAAGACTTCATCCAAAGCGCGAACG........GATAATGTATT  1116

194  eTrpGlyGluHisPheGluHisAspAsnLeuProAlaValArgAlaLeuA  211
     ||||||  ||::::::|||::||||||
1117 TTCGGGGCGAGAACTTTGAGTTCATGATGCTCCCAAAAATCGATGAAGTGT  1166
```

Fri Jan 19 06:20:05 2001

```
211  rgLeuHisLeuTyrArgAspSerAspLysLysArgLysLysAspLysAla  227
     ||:::||||||||||:|| //::|| ||||||||||||||||
1167 GCGTGAGCTTGTTCCAAGAATCTGATTCAAAGAAAAGAAGGACACT...  1213

228  GlyTyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgHi  244
     ||||||| ||:|| |||:::||||
1214 ...CTCATCGGATACGTTACCATTGGAATCGACCAACTGTCTAGCAGAAG  1260

244  sPheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGlyS  261
     ::|||||||| |||::  ::|||||:||||
1261 TCCAGTAGAGCGATGGTATACAGTCAACACATCTCATAGTGACTCTGGGA  1310

261  erGlyGlyMetGlySerGlyGlyGlyGlyGlyGlySerGlySerGly  277
     |||||||||||||| ||||||  //::||
1311 CAAGGAGAATTGCATCCGCACTAGGTGGGAAGAGCAGCTCCCAGGAATCC  1360

278  GlyLysGlyLysGlyGlyCysProAlaValAlaArgLeuLysAlaArgTyrGl  294
     |||||:::||||||||:::||||
1361         CCATCGCTTAGAATCAAAGCCCGTTGGCA  1389

294  nThrMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyrV  311
     ::||| ||||||||:|
1390 GTCGGTGCACATCCTCCCACTTCGAGCCTACGACAACCTTCTGGAAACAC  1439

311  aThrAsnHisTyrArgMetLeuCysAlaValLeuGluProAlaLeuAsn  327
     ::|||  ||||||||
1440 TTTGCTATAACTATTTGCCGCTTTGCGAGCAATTGGAGCCAGTGCTCAAT  1489

328  ValLysGlyValLysGluGluValAlaSerAlaLeuValHisIleLeuGlnSe  344
     |||| ||||
1490 GTCAGAGACAAGGAGGACTTGGCCACATCGTTGGTTCCTGTTATGTACAA  1539

344  rThrGlyLysAlaLysAspPheLeuSerAspMetAlaMetSerGluValA  361
     : ::|| ||||||| |||:::||||||
1540 ACACAACCTCGCAAAGGAGTTCCTGTGTGATTTGATCATGAGAGGAGTCG  1589

361  spArgPheMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAla  377
     ::||||::||||||:|||||| ||||||||||
1590 AGAAG...CTCGACAATGATCATTTAATGTTCAGAGGAAACACACTGGCC  1636

378  ThrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLy  394
     ||||||||:::||::||||::|||||||||
1637 ACAAAGGCTATGGAGTCGTTTATGAAACTTGTCGCCGACGATTATCTAGA  1686

394  sAspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsnC  411
     :|||||||:::||||:| ||:::||||:|
1687 CTCAACACTCAGTGATTTATTAAAACAGTGTTACAATGTGAGGGATTCAT  1736

411  ysGluValAspProIleLys.......CysThrAlaSerSerLeuAla  424
     ||||||||||| |||||||||||
1737 GCGAAGTAGATCCCACAGAAATTGGGTAATGTGTCAAACTCATCCCTCGAG  1786

425  GluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysVa  441
     ::|||||||| |||:||
1787 AAGAATCGTGCCCTTTGATGCGATATGTTGAAGTGGCTTGGACGAAAAT  1836

441  lValAsnSerHisCysValPheProArgGluLysLysGluValPheAlaS  458
     |||||:::||||||||:||||||||||||
1837 TTTGAACAACGTTCACCAGTACCCAAAAAACCTCCGACGTATTTTCGG  1886

458  erTrpLeuArgCysAlaGluGluArgGlyArgGluAspIleAlaAspArg  474
     :::||| ||| |||||
1887 GCGAAGTAGATGCCCGACTTGAAGCCCAGAATCGTGAAGCTTTGGCAGATACT  1936

475  LeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSe  491
     ||||::||||::||||||||||||||||:::||
1937 TTGATCTCCTCCATCAATCTTCCTTCGATTCTTATGTCCTGCAATTTTGAG  1986

491  rProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnThrSerA  508
     :::||| ||| |||:::||||||||||||||
1987 CCCAAGCCTTTTCAATCTTGTTTCCGAGTACCCATCACCAACCAACCGC  2036
```

```
508  rgThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSer  524
     ||:::|||||||||||||||||| :|||||||||||||||||||
2037 GCAACTTGACTCTGATTGCCAAGACGCTGCAGAACCTTGCGAATTTCAGC  2086

525  LysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGl  541
     ||||| ::: ||:|| |||||:::||
2087 AAGTTTGGAGGAAAAGAGCCTCACATGGAGTTTATGAACGAGTTTGTCGA  2136

541  uLeuGluTrpGlySerMetGlnGlnPheLeuTyrGluIleSerAsnLeuA  558
     ||||| ||:::||||||||: ::|||||||
2137 CCGAGAGTGGCATAGGATGAAAGATTTCTTGTTGAGAATTTCTTCAGAGT  2186

558  spThrLeuThrAsnSerSerPheGluGlyTyrIleAspLeuGlyArg  574
     ::|||||||||||||| ||||||||:::|
2187 CAAAGTCGGGACCAGAAAAAATGCGGATCAAATTGTGACGGCGGCCAAA  2236

575  GluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProGlnLeuSe  591
     ||||||| |||:::|| ::|||||||
2237 GAGCTTAGCTTAATCGCTACTTATCTTGAAGAAGCCTGGACTCCACTTCT  2286

591  rLysGluAlaLeuLeuLysLeuGlyProLeuProArg.......LeuL  605
     ::||| |||||| ::: ::::|
2287 CCAAGAGAAAATGGAAATAAGCATCCACTCTCAAATGTCAAGTCGGTTT  2336

605  euSerAspIleSerThrAlaLeuArg..................  613
     |||||| ||||||||
2337 TATCCGAATTGGCAGAATGCAAGAGACGCTCGGATAATGGCGTCTTTCAC  2386

614  AsnProAsnIleGlnArgGlnProSerArgGlnSerGluArgAlaArgSe  630
     ::||| ||||||| ::: :||||||||
2387 TCTCCAATGGTTCAA...CAGCCGTTCTTCAGATTATGAGAATAGTCCACA  2433

630  rGlnProMetValLeuArg................GlyP  638
     :: ::||| ||
2434 GCAACACGTTGTTCCTGCACATGAGAATGTACCGGCATATCGCAGTACTC  2483

638  roSerAlaGluMetGlnGlyTyrMetMetArgAspLeuAsnSerSerIle  654
     || ::::|| |||
2484 CGCCAACTGCCAAGCCACGGTATTGGGCCGTTCTTACAAACCGCCCT...  2530

655  AspLeuGlnSerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAl  671
     || ||:::||||||||||:||||||||||||||||
2531 ............GCAACTCATTTGCTCACATCGGACGATTACCT  2562

671  aArgLeuProSerProThrLysGluLysProProProProProGlyG  688
     ::: ||| ||||||||
2563 C......CTCTCTTCTGCATTTCAAACTCCAAGTCTTCGCCCCTGGAG  2603

688  lyGlyLysAspLeuPheTyrValSerArgProProLeuAlaArgSerSer  704
     ::||| |||||||||
2604 GC..................ACTCGGTTAAGC  2617

705  ProAlaTyrCysThrSerSerAspIleThrGluProGlnLysMe  721
     ||||||||||||||
2618 GATGAAACTGCCACGTCTTCGAGTCGCACCAGTGACAAGCAACCAGCAG  2667

721  tLeuSerValAsnLysSerVal...SerMetLeuAspLeuGlnGlyAsp.  736
     ::::||||||| ||| :::||:|| ||
2668 CGCTGGAGATTCGAGACGACACTGATTCGGATTTCGAGTTACGAGGGATC  2717

737  ..GlyProGlyArgLeuAsnSerSerValSerArgAsnLeuAlaAla  752
     :::||| ||| |||
2718 GAGGACGTGGGAGGAAGA...AACCGTAAGAGGCTACCACGTACTGATGCA  2764

753  ValGlyAspLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGl  769
     ||| ::::|| :| ||
2765 ..........TCACCATCGAGCAGTCAACAAGCTTCAAGTGG  2796

769  yLeuArgProAlaProAlaGlyArgLeuSerGlnGlyGlySerSerI  786
     :::||| ||| :::||||||||:|
2797 ATATTTAAGTAATAATCCTTCGAGATCACGTACTCAAACTTCTTCGAGTT  2846

786  leThrAlaAlaGlyMetArgLeuSerArgGlnMetGlyValThrAspGly  802
```

```
                                                    Ratio:      2.143                   Gaps:           39
                                     Percent Similarity:   57.859    Percent Identity:   36.788

                                     alignment_block:
                                     US-09-294-298-6 x AK022662

                                     Align seg 1/1  to: AK022662    from: 1  to: 2385

2847 CATCT...................................... 2851    454 GluValPheAlaSerTrpArgLeuArgCysAlaGlyArgGlyArgGluAs 470
     :::                                                      |||||||||||||||||:::  :  :::||||:::::||
803 ValProAlaGlnLeuArgIleProLeuSerPheGlnAsnProLeuPh 819          GAAGTGTTTGCATCATGGAAGCAGCAGTGCCTGAACCGTGCAACGCAAGA 51
     |||::::::::: |  |||||||| |||||| ||
2852 ..CCAGTTGAACGAATG...GCCGCCTCTATCAATTGCTAACCCATGCTT 2895  2 GAAGTGTTTGCATCATGGAAGCAGCAGTGCCTGAACCGTGCAACGCAAGA 51

819 eHisMetAlaAlaAspGlyProGlyProGlyProAlaGlyHisGly....... 833  470 pIleAlaAspArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysP 487
     |                                                          |||::::::||||||||||||:::::||||||||||||||
2896 T...............GGACCAGGCCCATCATCTGGATATGCTATACCTG 2930      CATCAGCGAGAGGCTCATCAGTGCCCTCATTATTTCCCGTTTCTGTGTC 101

834 ..............GlySerSerGlyHisGlyProProSer 842            487 roAlaIleMetSerProSerLeuPheGlyLeuMetGlnGluTyrProAsp 503
     :::          |||||||                                      ||||||||||||||||||||||||||||||||||||||||
2931 CAGAGCCAAAGGAAATCGTATACACAAAAGCGAGCAAGCCCACCACCATAC 2980   102 CAGCCATTATGTCTCCCAGTCTTTTCAACCTTATGCAGGAGTATCCTGAT 151

843 SerHisHisHisHisHisHis 851                               504 GluGlnThrSerArgThrLeuThrLeuIleAlaLysValIleGlnAsnLe 520
     |||:::::::::::||||:                                       :::::||||||||||||:::||||||||||||||||
2981 GATCCCGATGTGCACAACTATCATTAT 3007                        152 GACCGGCACATCTCGGACTCTAACTCTTATTGCCAAGGTCATTCAGAACCT 201

seq_name: gb_pr4:AK022662                                   520 uAlaAsnPheSerLysPheThrSerLysGluAspPheLeuGlyPheMetA 537
                                                                 ||||||||::::||||:::||||||
seq_documentation_block:                                     202 GGCCAACTTTGCCAAGTTTGGTAACAAAGAGGAATACATGGCATTCATGA 251
LOCUS       AK022662    2385 bp    mRNA          PRI   29-SEP-2000
DEFINITION  Homo sapiens cDNA FLJ12600 fis, clone NT2RM4001412, highly similar  537 snGluPheLeuGluLeuGlyTrpGlySerMetGlnGlnPheLeuTyrGlu 553
            to Homo sapiens nGAP mRNA.                           ||::::||||||    ||||||||||||:::||||||
ACCESSION   AK022662                                        252 ATGATTTTTAGAACATGAATGGGGTGGAATGAAGCCGCTTTCTTTGGAG 301
VERSION     AK022662.1   GI:10434186
KEYWORDS    oligo capping; fis (full insert sequence).      554 IleSerAsnLeuAspThrLeuThrAsnSerSerPheGluGlyTyrTyrIl 570
SOURCE      Homo sapiens teratocarcinoma cell line:NT2 cDNA to mRNA,  ||||||||||||
            clone_lib:NT2RM4 clone:NT2RM4001412.            302 ATCTCTAATCCAGACACCATCTCAAACACCCCAGGCTTTGATGGTTACAT 351
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  570 eAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValL 587
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  |||||||||||:::|||||||
REFERENCE   1 (sites)                                       352 TGATCTGGGCCGAGAGCTTTCAGTTTGCATTCCTACTGTGGGAAGTAG 401
AUTHORS     Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
            Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,  587 euProGlnLeuSerLys................GluAlaLeuLeu 596
            Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,  :: ||||||::||| |
            Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,  402 TTTCCAACTTGATAAGGTGAAAATTCCTTCCTACAGGCGACCGTGGCA 451
            Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
            Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,  597 LysLeuGlyProLeuProArgLeuLeuSerAspIleSerThrAlaLeuAr 613
            Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.  |||||||||||||||   :::::::::
TITLE       NEDO human cDNA sequencing project               452 AAATTGGGGCCTCTCCCCGTGTTCTTGCTGTGATATTACCAAGTCATTGAC 501
JOURNAL     Unpublished (2000)
REFERENCE   2 (bases 1 to 2385)                             613 gAsnProAsn...IleGlnArgGlnSerArgGlnSerGluArgAlaA 629
AUTHORS     Isogai,T. and Otsuki,T.                              ||||||:::   ||||||:::|||
TITLE       Direct Submission                               502 TAATCCTACGCCAATACAACAGCAACTGAGACGCGTTCACTGAACAT.... 547
JOURNAL     Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
            Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,  629 rgSerGlnProMetValLeuArgGlyProSerAlaGluMetGlnGlyTyr 645
            Kisarau, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,  ||||||||||||||||||||||||||||
            Tel:81-438-52-3951, Fax:81-438-52-3952)         502 TAAAATCTTCCAGCCAGGACAAC......ACAGACAGCTACTTCAGAGGG 676
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert  547 ................................................ 547
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:  646 MetMetArgAspLeuAsnSerIleAspLeuGlnSerPheMetAlaAr 662
            Helix Research Institute (supported by Japan Key Technology Center  ||||||| ::::::::::::::::::::
            etc.) and Department of Virology, Institute of Medical Science,  548 ...........AACTCCAGTCCAAATGTCAGTGGAAGCCCTTCCTC 582
            University of Tokyo.
FEATURES    Location/Qualifiers                             662 gGlyLeuAsn.................SerSerMetAspMetAlaArgL 673
     source  1..2385                                             ||:::::                   ||:||::   :::||
            /organism="Homo sapiens"                        583 TGGGCTGCAGAAAATATTTGAAGACCCCACCGATCAGTGATTTGCATAAAC 632
            /db_xref="taxon:9606"
            /cell_line="NT2"                                673 euProSerProThrLysGluLysProProProProProGlyGlyGly 689
            /cell_type="teratocarcinoma"                        || |||||||||||||||
            /clone="NT2RM4001412"                           633 TAAAATCTTCCAGCCAGGACAAC......ACAGACAGCTACTTCAGAGGG 676
            /clone_lib="NT2RM4"
            /note="cloning vector: pME18SFL3-mRNA from uninduced NT2  690 LysAspLeuPheTyrValSerArgProLeuAlaArgSerSerProAl 706
            neuronal precursor cells."                          |||  ||||::: ||:::::||
BASE COUNT  690 a   602 c   576 g   517 t                   677 AAACATTATTGCTGGTTCAGCAA.........GCCTCCTCTCAGAG 714
ORIGIN
                                                            706 aTyrCysThrSerSerSerAspIleThrGluProGluGlnLysMetLeuS 723
alignment_scores:
     Quality: 1088.50      Length:      878
```

Fri Jan 19 06:20:05 2001

```
1390  GAATGGGAGCCGGTCCGGCAGCAGTCCTTCCTCCTCCAGAGAGAGCCCTG  1439
      leGlyGlySerGlyGlySerGlyGlyGlyGlyGlyGlyGlyLeuLysPro
1002  ::                                             |||  1018
1440  TT.........................................CCC  1444
      SerIleThrLys...GlnHisSerGlnThrProSerThrLeuAsnPro..
1019  :::|||       |||:::::::::|||:  ||::||||||||||||||  1033
1445  AAAGTTAGAGCAATCCAGAGACAACAGACACAGCAGGTTCAGTCACCTGT  1494
               ..ThrMetProAlaSerGluArgThrValAlaTrpValSerA
1034  ||||||               |||||||||||:::||||||||  |     1047
1495  GGACTCTGCCACAATGTCCCCAGTAGAGAGACAGCAGCCTGGGTTCTGA  1544
      snMetProHisLeuSerAlaAspIleSerAlaHisIleGluArgGlu
1047  |||::||                                             1063
1545  ACAATGGCAGTATGAAGAGATGTGGAA.....................   1573
      GluTyrLysLeuLysGlyTyrSerLysSerMetAspGluSerArgLeuAs
1064  :::|||||::                                          1080
1574  .........GAAACTGAGCAAAATCTAGATGAAGCCAAG........   1603
      pArgValLysGluTyrGluGluGlnIleHisSerLeuLysGluArgLeuH
1080  :::::::::::::|||:::|||||||:|||||||||||||||||        1097
1604  .CATGCTGAGAGTATGAACAAGAAATTACTAAACTGAAGGAGCGCCTGA  1652
      isMetSerAsnArgLysLeuGlyGluTyrGluArgArgLeuLeuSerGln
1097  |||||:::|||::||||||||||||||||||||||||||||||||||     1113
1653  GAGTTTCCACCCGGCGACTGGAGAATATGAACGCCGCTTGCTGGTGCAG  1702
      GluGluGlnThrSerLysIleLeuMetGlnTyrGlnAlaArgLeuGluGl
1114  |||::|||    :|||::|||||:::|||::||||||||||||||:::    1130
1703  GAGCAGCAGATGCAGAAGCTGCTGCTGGAATACAAGGCCCGACTGGAGGA  1752
      nSerGluLysArgLeuArgGlnGlnGlnValGluLysAspSerGlnIleL
1130  |||:::|||    |||||||||:||  ||||||||||||::||||||||   1147
1753  CAGCGAGGAGCGGCTCCGAGACAGCAGGAGAAGAAAAGATAGCCAGATGA  1802
      ysSerIleIleGlyArgLeuMetLeuValGluGluValLeuArgArgAsp
1147  |||||||||:::|||||||||  |||||||||||||||:::|||||||||  1163
1803  AAAGCAATCATCAGCAGGCTAATGGCTGTGGAAGAGGAACTGAAGAAGGAT  1852
      HisProAlaMetAlaGluMetAlaGluProGluProLysLysArgLeuAs
1164  |||   :::                ||:||| ||||||||||||||||||  1180
1853  CATGCTGAGATG...CAAGCAGTTATTGATGCAAAGCAGAAAATAATTGA  1899
      pAlaGln......................Arg..GlySe
1180  ||||||                  ::||||    ::||              1185
1900  TGCACAGGTCATAAATGGAGATGAGATTATTCAAACAGGAAAAACGGATC  1949
      rPheProProTrpValGlnGlnThrArg........ValArgTrpPro.
1185  ||||||||::|||  |||                       ::|||  |   1198
1950  GTGT...CCCTGGATTCAGCCAACACCAGACTGATGAGGCCGCTGACCCA  1996
      .....HisLeuGlyThrAlaTrp...................ProProGln
1199  :::  ||||||||||                          ||||||||| 1207
1997  AGTGAAGGAGCGGTACACATGCAGGTCCGCAATGGCATCTCCCCCACCA  2046
      ProHisProProHisProGlyCysArg........SerGlnArgThrAl
1208  |||||||||||  |||:::||                ||||  :::||    1221
2047  ACCCCACCAAGCTTCCATCACGGAGAATGGTGAATTCAAAAACAGCAGC  2096
      aSerSerGlyThr.......................ProGlnT
1221  |||||||||:::|||                      |||            1228
2097  TGCTGACAGGCGCTTTGTCTGTGGAAGGAGACAGAGGAAATTGACCACTC  2146
      hrThrSerProProSerIleThrAspLeuLeuProCysAlaProTyrPro  1244
1228  |||||||                        |||||               
2147  TCCTATCTCCA.........................GACCTT.......   2169
      GlyProProSerValThrAspLeuLeuProSer  1255
1245  ::|||||:::::  :::::  |||                            
2170  GCCCCTCCAGGTTTACAGAATGTTGCTACTTCA  2202
```

```
715  :                     ||:::::|||                 755
     CATGACTTATTCTGAAAAGGAT..........GAAAGGAAAGTAGCCTTC
     ervalAsnLysSerValSerMetLeuAspLeuGlnGlyAspGlyProGly
723  |||||||||:::||||||||:::|||||||                    739
756  CTAATGGTCGGGAGCGTCTCCCTCATGGACCTCCAG..GACACTCATGCT  802
     GlyArgLeuAsnSerSerSerValSerAsnLeuAlaAlaValAlaGlyAspLe
740  :::|||||||:::||||||         :::                    756
803  GCTCAAGTGGAGCATGCATCTGTC......ATGCTTGATGTGCCTATACG  846
     uLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProA
756  |||  :::||||||  |||:::||  :::::::::                773
847  CTTGACCGGAAGCACCAGCCTTTCCATAACCCAGGTGGCCAGCATCAAA...  892
     laProAlaGlyLeuArgGlnGlySerGlySerSerIleThrAlaAla
773  :::                         ||:::|||              789
893  .........CAGCTGCGGAAACCCAGAGCACTCCCCAAAGTGCACCC      931
     GlyMetArg...............LeuSerGlnMetGlyValThrTh
790  ::|||                  |||:::|||  |||              800
932  CAAGTGAGAGAAGGGCCCCTGCACCCACGCTTGAACCAGCCAGT........  973
     rAspGlyValProAlaGlnGlnLeuArgIleProLeuSerPheGlnAsnP
800  |||:::                  ||:::|||||||||||||         817
974  ...GGCCTTCAG...................CCCTTGTCGTTCCAGAACC  1001
     roLeuPheHisMetAlaAlaAlaAspGlyProGlyProProAlaGlyHisGly
817  |||:::||:::||||||                                  833
1002 CTGTCTATCACCTCAATAACCAATTCCAGCAATGCCAAAGGCCTTCTATA  1051
     GlySerSerGlyHisGlyProProSerSerHisHisHisHisHisHi
834  |||::||||||||                   |||               850
1052 GATTCCAGT.......................................   1060
     sHisHisArgGlyGlyGluProProGlyAspThrPheAlaProPheH
850  :::|||||                                          867
1061 ...............................TTGGAGAACCTAAGCACT  1078
     isGlyTyrSerLysSerGluAspLeuSerThrGlyValProLysProPro
867  :::|||||||||   |||  :::||||||||||:::|||            883
1079 GCCAGTTCC.....AGAAGCCAAAGTAACAGTGAAGACTTCAAGCTCAG  1122
     AlaAlaSerIleLeuHisSerHisSerTyrSerAspGluPhe........
884  |||||||||                          :::            897
898  .GlyProSerGlyThr.........AspPheThrArgArgLeuLeuSerL
911  ||||||||:::||                 ||||||:::|||:::||    
    TGGACCCCAGCAATAGCAGCAGCAGGAAGATTTCACTAAACGTAGCACTCAGA  1172
1123 |||||||                              :::          
    .GlyLysSerGlyAsnLeuLeuGlnLeuGlyLeuIleSerProSerGlyI  986
958  lyLysSerGlnLeuThr.........ValSerAlaAlaGln  969
    GTGCTTCTTTACGTAGCACGGGAGCACTGTCAGTGGTCCCGCAGCC...  1369
1323 |||||||||||                      |||||||||||      
970  LysProArgProSerSerGlyAsnLeuGlnLeuGlnSerProGlySerTy  986
    .........................CTGGTGGCCGAACCTGTGCA  1389
1370                          |||::::||:::||||:::
986  rGlyProAlaAlaArgProArgGlnGln...SerLeuSerGluLysGlySerI  1002
    :::|||  ||||||||||  |||:::||||||
```

```
OM of: US-09-294-298-6 to: EST:*    out_format : pfs

Date: Jan 18, 2001  9:24 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09294298/runat_17012001_125501_27745/app-query.fasta_1.4359
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09294298_@CGN1_1_7119 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXV -WAIT -THREADS=1

Search information block:
Query: US-09-294-298-6
Query length: 1325
Database: EST:*
Database sequences: 7991742
Database length: -791223438
Search time (sec): 4688.750000
```

sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est37:AV590698	+	933.00	1006.68	2.1e-50	571	AV590698 AV590698 Bos taurus br
gb_gss21:AZ399131	-	764.00	874.89	1.0e-39	474	AZ399131 1MO164P14R Mouse 10kb
gb_est49:AW779747	-	737.50	841.95	6.9e-38	593	AW779747 hn85f10.x1 NCI_CGAP_Ki
gb_est44:AW421227	-	731.50	834.99	1.7e-37	597	AW421227 fj94f10.x1 zebrafish g
gb_est71:BE490923	-	694.50	791.97	4.2e-35	628	BE490923 db38g08.x1 Blackshear/
gb_est22:AI606711	+	660.00	753.57	5.8e-33	566	AI606711 m158d01.y1 Stratagene
gb_est71:BE463433	-	628.50	719.70	4.5e-31	463	BE463433 hw23d06.x1 NCI_CGAP_Ki
gb_est22:AI577567	-	600.00	685.89	3.4e-29	511	AI577567 UI-R-Y0-vk-d-04-0-UI.s
gb_est41:AW205989	+	571.50	654.04	2.0e-27	474	AW205989 UI-H-BI1-afy-d-12-0-UI
gb_est28:AJ396547	+	566.50	643.77	7.6e-27	707	AJ396547 dkfz426 Gallu
gb_est11:AA709010	-	563.00	645.47	6.1e-27	427	AA709010 zf94h11.s1 Soares_pine
gb_est22:AI609604	-	548.50	627.99	5.7e-26	460	AI609604 tw91e09.x1 NCI_CGAP_HN
gb_est9:AA554056	-	543.50	622.53	1.2e-25	449	AA554056 nl01g07.s1 NCI_CGAP_Co
gb_est27:AI958829	+	537.50	613.24	3.8e-25	555	AI958829 fd22q01.y1 zebrafish W
gb_est40:AW136165	-	531.50	608.91	6.6e-25	443	AW136165 UI-H-BI1'-abz-d-09-0-UI
gb_est9:AA554055	-	517.50	592.62	5.3e-24	452	AA554055 nl01g06.s1 NCI_CGAP_Co
gb_est10:AA691827	+	504.00	578.16	3.4e-23	413	AA691827 vt05d01.r1 Barstead mo
gb_gss26:CNS02VGW	+	501.00	564.47	2.0e-22	1020	AL215753 Tetraodon nigroviridi
gb_est74:BE695645	-	486.00	554.98	6.7e-22	516	BE695645 MR1-BT0800-030700-002-
gb_est9:AA575948	-	484.50	556.76	5.3e-22	379	AA575948 nm56e05.s1 NCI_CGAP_Br
gb_est22:AI562171	-	484.50	553.74	5.4e-22	382	AI562171 vw73f04.x1 Stratagene
gb_est41:AW204612	+	481.00	553.74	7.8e-22	347	AW204612 UI-H-BI1-aej-c-02-0-UI
gb_est2:AA107246	+	480.00	549.90	1.3e-21	440	AA107246 m158d01.r1 Stratagene
gb_est39:AW076911	+	473.00	544.63	2.5e-21	345	AW076911 fj03d08.y1 Zebrafish a
gb_est76:BE848541	-	464.00	528.11	2.1e-20	596	BE848541 uw39f05.y1 Soares_thym
gb_gss26:CNS02JDV	+	461.50	519.62	6.2e-20	979	AL200092 Tetraodon nigroviridis
gb_gss26:CNS02JKG	+	457.50	516.87	8.9e-20	832	AL200329 Tetraodon nigroviridis
gb_gss14:AQ993964	+	454.00	515.11	1.1e-19	682	AQ993964 RPCI-23-241C10.TV RPCI
gb_gss26:CNS01XQ5	-	452.50	514.27	1.3e-19	631	AL172022 Tetraodon nigroviridis
gb_gss25:CNS00HWF	-	425.50	478.01	1.3e-17	1006	AL073826 Drosophila melanogast
gb_gss26:CNS03ANH	+	419.50	471.48	3.0e-17	975	AL235430 Tetraodon nigroviridis
gb_est45:AW527364	+	418.50	483.06	6.8e-18	317	AW527364 UI-R-BO1-ajj-e-01-0-UI
gb_est59:BB321064	+	417.50	480.38	9.6e-18	363	BB321064 BB321064 RIKEN full-le
gb_est26:AI874961	+	416.50	480.11	9.9e-18	336	AI874961 u127f03.x1 Sugano mous
gb_gss27:CNS04DL0	+	416.00	469.24	4.0e-17	877	AL285885 Tetraodon nigroviridis
gb_est4:AA258366	-	401.50	463.49	8.3e-17	319	AA258366 zr59e01.s1 Soares_NhHM
gb_gss27:CNS041WW	+	401.00	450.84	4.2e-16	926	AL270761 Tetraodon nigroviridis
gb_est61:BB394547	+	394.00	448.87	5.4e-16	542	AW656799 109245 MARC 1BOV Bos t
gb_est23:AI650331	+	383.50	442.62	1.2e-15	325	BB394547 BB394547 RIKEN full-le
gb_est14:AA953060	-	379.00	438.52	2.1e-15	296	AI650331 wa18f01.x1 NCI_CGAP_Ki
	-	368.50	426.63	9.4e-15	292	AA953060 oo01g09.s1 Soares_NFL_

```
gb_est59:BB312776                  + 365.50  422.54  1.6e-14  309 ! BB312776 BB312776 RIKEN full
gb_est75:BE773015                  - 365.50  417.53  3.0e-14  481 ! BE773015 RCI-FT0134-170700-0
gb_est19:AI327335                  - 358.00  405.50  1.4e-15  651 ! AI327335 mp74c11.x1 Soares_t
gb_est66:BE085114                  - 357.50  412.45  5.8e-14  335 ! BE085114 CM2-BT0664-240200-1

seq_name: gb_est37:AV590698

seq_documentation_block:
LOCUS       AV590698           571 bp    mRNA            EST    29-AUG-2000
DEFINITION  AV590698 Bos taurus brain fetus Bos taurus cDNA clone E1BR014B04
            5', mRNA sequence.
ACCESSION   AV590698
VERSION     AV590698.1  GI:9701691
KEYWORDS    EST.
SOURCE      cow.
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 571)
AUTHORS     Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
            Suzuki.H.
TITLE       bovine cDNA sequencing
JOURNAL     Unpublished (2000)
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES             Location/Qualifiers
     source          1..571
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="E1BR014B04"
                     /clone_lib="Bos taurus brain fetus"
                     /tissue_type="brain"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
                     was deleted from a NotI site"
BASE COUNT       124 a      161 c      169 g      116 t      1 others
ORIGIN

alignment_scores:
              Quality:     933.00             Length:      186
              Ratio:         5.043             Gaps:          1
Percent Similarity:  99.462    Percent Identity:  99.462

alignment_block:
US-09-294-298-6 x AV590698

Align seg 1/1  to: AV590698   ..  from: 1  to: 571

  331 LysGluGluValAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLyLy 347
      ||||||||||||||||||||||||||||||||||||||||||||||||||||
   13 AAGGAGGAGGTCGCCAGTGCTCTCGTTCACATCCTGCAGAGTACAGGCAA    62

  347 sAlaLysAspPheLeuSerAspMetAlaMetSerGluValAspArgPheM  364
      |||||||||||||||||||||||||||||||||||||||||||||||||||
   63 GGCCAAGGACTTCCTTCAGACATGGCCATGTCCGAGGTGGACCGGTTCA   112

  364 etGluArgGluHisGluPheArgGluAsnThrLeuAlaThrLysAla  380
      |||||||||||||||||||||||||||||||||||||||||||||||
  113 TGGAACGAGAACACCTCATATTCCGCGAGAACACGCTCGCCACTAAAGCC   162

  381 IleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLysAspAlaIl  397
      |||||||||||||||||||||||||||||||||||||||||||||||||
  163 ATAGAAGAGTACATGAGACTGATTGGTCAGAAATACCTCAAGGATGCCAT   212
```

```
397 eGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsnCysGluValA 414
    ||||||||||||||||||||||||||||||||||||||||||||||
213 CGGGGAATTCATCCCGTGCTCTGTATGAATCTGAGGAGAACTGTGAGGTGG 262

414 spProIleLysCysThrAlaSerSerLeuAlaGlyGluHisGlnAlaAsnLeu 430
    ||||||||||||||||||||||||||||||||||||||||||||||
263 ACCCCATCAAGTGCACGGCATCCAGTTTGCTGAGCACCAGGCCAACCTG 312

431 ArgMetCysCysGluLeuAlaLeuCysLysValValAsnSerHisCysVa 447
    ||||||||||||||||||||||||||||||||||||||||||||||
313 CGGATGTGCTGTGAGTTGGCCCTGTGCAAGGTGGTCAACTCCCATTGCGT 362

447 lPheProArgGluLeuLysGluValPheAlaSerTrpArgLeuArgCysA 464
    ||||||||||||||||||||||||||||||||||||||||||||||
363 GTTCCCGAGGGAGCTGAAGGAGGTGTTTGCGTCCTGGCCACTCGCGTCTG 412

464 laGluArgGlyArgGluAspIleAlaAspArgLeuIleSerAlaSerLeu 480
    ||||||||||||||||||||||||||||||||||||||||||||||
413 CAGAGCGGGGCCGAGAGACATTGCGACCGGCTGATCAGCGGCCTCGCTC 462

481 PheLeuArgPheLeuCysProAlaIleMetSerProSerLeuPhe.GlyL 497
    ||||||||||||||||||||||||||||||||||||||  .|||
463 TTCCTGCGCTTCCTCCTGCCCGCCCATTATGTCGCCCAGCCTNTTCGGGGC 512

497 euMetGlnGluTyrProAspGlnGlnThrSerArgThrLeuThrLeuIle 513
    ||||||||||||||||||||||||||||||||||||||||||||||
513 TCATGCAGGAGTACCCAGATGAGCAAACTTCGAGAACCCTCACGCTCATC 562

514 AlaLys 515
    |||||
563 GCCAAG 568
```

seq_name: gb_gss21:AZ399131

```
seq_documentation_block:
LOCUS       AZ399131              474 bp    DNA         GSS       03-OCT-2000
DEFINITION  1M0164P14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0164P14 R, DNA sequence.
ACCESSION   AZ399131
VERSION     AZ399131.1  GI:10514203
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 474)
AUTHORS     Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
TITLE       Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL     Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0164  row: P  column: 14
            Seq_primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 474.
FEATURES             Location/Qualifiers
     source          1..474
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0164P14"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="vector: PWD42nv: Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained  from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1). a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT       92 a      151 c     131 g    100 t
ORIGIN

alignment_scores:
        Quality:       764.00          Length:      142
        Ratio:          5.380          Gaps:          0
Percent Similarity:    100.000     Percent Identity:  99.296

alignment_block:
US-09-294-298-6 x AZ399131/rev  ..
Align seg 1/1  to reverse of: AZ399131  from: 1  to: 474

141 LysAspAsnSerArgArgValAspAsnValLeuLysLeuTrpIleIleGl 157
    ::|||||||||||||||||||||||||||||||||||||||||||||
426 CAGGACAACAGCCGCCGAGTAGATAAGCTGTCGAAGCTATGGATCATAGA 377

157 uAlaArgGluLeuProProLysArgTyrTyrCysGluLeuCysLeuA 174
    ||||||||||||||||||||||||||||||||||||||||||||||
376 GGCTCGAGAGCTGCCCCCAGAGAGAGATATTACTGTGAGCTGTGCCTGG 327

174 spAspMetLeuTyrAlaArgThrThrSerLysProArgSerAlaSerGly 190
    ||||||||||||||||||||||||||||||||||||||||||||||
326 ACGACATGCTGTATGCACGAACCACCTCCAAGCCCCGCTCGGCTTCAGGA 277

191 AspThrValPheTrpGlyGluHisPheGluPheAsnAsnLeuProAlaVa 207
    ||||||||||||||||||||||||||||||||||||||||||||||
276 GACACCGTCTTTGGGGCGAGCACTTTGAGTTTAACACCTGCCTGCCGT 227

207 lArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysArgLysL 224
    ||||||||||||||||||||||||||||||||||||||||||||||
226 CCGGGCCCTTCGGCTGCATCTGACCGTGACTCAGACAAAAAGCGGAGAA 177

224 ysAspLysAlaGlyTyrValGlyLeuValThrValProValAlaThrLeu 240
    ||||||||||||||||||||||||||||||||||||||||||||||
176 AGGACAAGGCTGGCTACGTGGCCTGGCACGTGTTCCAGTGGCCACCCTA 127

241 AlaGlyArgHisPheThrGluInTrpTyrProValThrLeuProThrGl 257
    ||||||||||||||||||||||||||||||||||||||||||||||
126 GCTGGGCGCCACTTCACAGAGCAGTGGTACCCCGTGACCTTGCCGACAGG 77

257 ySerGlyGlySerGlyGlyMetGlySerGlyGlyGlyGlySerGlyG 274
    ||||||||||||||||||||||||||||||||||||||||||||||
76  CAGTGGGGGCTCTGGGGGACCCGGGGAGGGGGGAGGAGGGAGGTCAGGGG 27

274 lyGlyGlySerGlyGlyLysGlyGly 282
    ||||||||||||||||||||||||||
26  GTGGCTCAGGGGGCAAAGGGAAAGGA 1
```

seq_name: gb_est49:AW779747

```
285 ProAlaValArgLeuLysAlaArgTyrGlnThrMetSerIleLeuProMe 301
    ||| ::||||||||||  :::|||||||||||||||:::||||||
503 CCCATGATCAGCATCAAGGGCGGCTACCAACCATCACCATCCTGCCCAT 454

301 tGluLeuTyrLysGluPheAlaGluTyrValThrArgAsnHisTyrArgMetL 318
    |||||||||||||||||||||||||||||||||||||
453 GGAGATGTACAAAGAGTTCGCTGAGCAATCACCAACCACTACCTGTGGC 404

318 euCysAlaValLeuGluProAlaLeuAsnValLysGlyLysGluGluVal 334
    |||||||| ||:::|||||  |||:::|||| ||||||||:::||||||
403 TGTGTGCAGCCCTGCAGCCCATCCTCAGTGCCAAGACCAAGGAGGAGATG 354

335 AlaSerAlaLeuValHisIleLeuGlnSerThrGlyLysValAlaLysAspPh 351
    |||||||||||||||||||||||||||||||:::||||||
353 GCATCTGCCCTGGTGCACATCCTGCAGAGCACGGGCAAGGTGAAGGACTT 304

351 eLeuSerAspMetAlaMetSerGluValAspArgPheMetGluArgGluH 368
    ||||:::||||  :::|||||||||||||
303 CCTGACAGACCTGATGATGTCAGAGGTGGACCGTGCGGGGACAACGAGC 254

368 isLeuIlePheArgGluAsnThrLeuAlaThrLysAlaIleGluGluTyr 384
    |||||||||||||||||||||||||||||||||||
253 ACCTCATCTTCCGGGAGAACACACTGGCCACCAAGGCCATTGAGGAGTAC 204

385 MetArgLeuIleGlyGlnLysTyrLeuLysAspAlaIleGlyGluPheIl 401
    :::||||||||||||  :::|||||||||||||||||
203 CTCAAGCTAGTGGGCCAGAGTACNTGCAGGACGGCCCTAGGTGAGTTCAT 154

401 eArgAlaLeuTyrGluSerGluGluAsnCysGluValAspProIleLysC 418
    |:::|||||||||||||  :::|||||||||||||||||
153 CAAAGCGCTGTATGAGTCAGATGAGAACTGCGAAGTGGATCCCAGCAAGT 104

418 ysThrAlaSerSerLeuAlaGluHisGlnAlaAsnLeuArgMetCysCys 434
    |||:::||||  :::|||||||||||||||||||
103 GCTCCGGCCGCTGACCTCCCAGACACCAGGCAACCTCAAGATGTGCTGC 54

435 GluLeuAlaLeuCysLysLysValValAlaAsnSerHisCysValPheProArgGl 451
    :::||||||||||  :::|||||||||||||||||||
53 GAGCTGGCCTTGTGCAAGATCATCAACTCCTACTGTGTCTTCCCACGGGA 4

451 u 451
3 G 3
```

seq_name: gb_est44:AW421227

seq_documentation_block:
LOCUS   AW421227   597 bp   mRNA   EST   09-FEB-2000
DEFINITION fj94f10.x1 zebrafish gridded kidney Danio rerio cDNA 3' similar to
            TR:O95174 O95174 NGAP. ;, mRNA sequence.
ACCESSION   AW421227
VERSION   AW421227.1   GI:6949159
KEYWORDS   EST.
SOURCE   zebrafish.
ORGANISM   Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Rasborinae; Danio.
REFERENCE   1 (bases 1 to 597)
AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
TITLE   WashU Zebrafish EST Project 1998
JOURNAL   Unpublished (1998)
COMMENT   Other_ESTs: fj94f10.y1
            Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810

seq_documentation_block:
LOCUS   AW779747   593 bp   mRNA   EST   12-MAY-2000
DEFINITION hn85f10.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3034699 3'
            similar to TR:O95174 O95174 NGAP. ;, mRNA sequence.
ACCESSION   AW779747
VERSION   AW779747.1   GI:7794350
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 593)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 451.
FEATURES   Location/Qualifiers
source   1..593
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3034699"
            /clone_lib="NCI_CGAP_Kid11"
            /lab_host="DH10B"
            /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
            a modified polylinker; Site_1: Not I; Site_2: Eco RI;
            Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
            prepared, and ss circles were made in vitro. Following HAP
            purification, this DNA was used as tracer in a subtractive
            hybridization reaction. The driver was PCR-amplified cDNAs
            from a pool of 5,000 clones made from the same library
            (cloneIDs 1322376-1323911, 1456007-1456775, and
            1500552-1502855)..Subtraction by Bento Soares and M.
            Fatima Bonaldo. "
BASE COUNT   104 a   169 c   178 g   141 t   1 others
ORIGIN

alignment_scores:
    Quality: 737.50   Length: 217
    Ratio: 4.143   Gaps: 2
Percent Similarity: 82.028   Percent Identity: 63.594

alignment_block:
US-09-294-298-6 x AW779747/rev ..

Align seg 1/1   to reverse of: AW779747   from: 1 to: 593

```
104 oValThrLeuProThrGlySerGlyGlySerGlyGlyMetGlySerGlyG 268
    ||||  |||:|||   ||:||||   ::||||
543 GGTGTTGACGCCCAACCCCAGGGC............................. 519

236 ProValAlaThrLeuAlaGly...ArgHisPheThrGluGlnTrpTyrPr 251
    |||:::::: |||:::|||  |||||||||||:::|||||
593 CCTACCTGCTGCCTCGGTGCCGCGGCCAGTTCGTGAGAAGTGGTACCC 544

268 lyGlyGlyGlySerGlyGlySerGlyGlyLysGlyLysGlyGlyGlyCys 284
    ||||||||
518 .........................GGCAAGGGCACTGGA 504
```

```
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Leonard Ira Zon DNA Sequencing by:
            Washington University Genome Sequencing Center Clone distribution:
            Genome Systems, St. Louis, Missouri (web address:
            www.genomesystems.com) (email contact: info@genomesystems.com) and
            Research Genetics, Huntsville, Alabama (web address: www.resgen.com
            ) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Possible reversed clone: polyT not found
            Seq primer: T7 ET from Amersham
            High quality sequence stop: 409.
  FEATURES              Location/Qualifiers
       source          1..597
                       /organism="Danio rerio"
                       /db_xref="taxon:7955"
                       /clone_lib="zebrafish gridded kidney"
                       /sex="mixed"
                       /tissue_type="kidney pooled from 300 wild type adults"
                       /lab_host="XLOLR"
                       /note="Organ: kidney; Vector: pBK-CMV; Site_1: EcoRI;
                       Site_2: XhoI; Oligo dT cDNA library constructed from mRNA
                       pooled from pooled kidney tissue from 300 adult
                       zebrafish."
  BASE COUNT    132 a    153 c    150 g    162 t
  ORIGIN

  alignment_scores:
       Quality:     731.50          Length:      202
       Ratio:         4.041          Gaps:          2
  Percent Similarity:  89.604     Percent Identity:  67.327

  alignment_block:
  US-09-294-298-6 x AW421227/rev

  Align seg 1/1 to reverse of: AW421227    from: 1  to: 597

   18 GlyPheLeuSerArgArgLeuLysSerSerIleLysArgThrLysSerGl  34
      |||||||||||||||:::||||||    ::||||||||||||||||||||
  591 GGTTTTTTAAGTAAACGACTGATAGGATCCATCAAAAGGACGAAGAGCCA  542

   34 nProLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheA  51
      |  |||||||||||||||||||    :||||| ||||  :||||
  541 GACCAAGCTGGACCGCAACACAAGTTCAGA.......CTCCCATCGGTGC  498

   51 rgSerAlaAspHisAspArgAlaArgLeuMetGlnSerPheLysGluSer  67
      ||  :::||||||||||||||::|||   :::  ::|||||||
  497 GACCAACAGAAAATGACAGGTCACGAGGCCTGCCCAAATTGAAGGAGTCC  448

   68 HisSerHisGluSerLeuLeuSerProSerSerAlaAlaGluAlaLeuGl  84
      ||||||||||||||||||::|||||:::||||||  :::||||||||||
  447 TGTTCCCATGAGTCTTTGCTGAGTCCTGGCAGCGCTGTTGAAGCTCTGGA  398

   84 uLeuAsnLeuAspGluAspSerIleLeuLysProValHisSerSerIleL  101
      :|||:::::::||||||    :::  |||||||||||||||||||
  397 TCTGAGTATGGAGGAGGATGTCTACATCAAACCTTTACACAGCAGCATCC  348

  101 euGlyGlnGluPheCysPheGluValThrThrSerSerGlyThrLysCys  117
      ||||||||||||||||||||||::  ||::::||||::::|||||
  347 TAGGACAAGATTTCTGCCTTTGAGTGGCGTACTCAGGTGGAAGTAAGTGC  298

  118 PheAlaCysArgSerAlaAlaGluAlaGlyArgAspAlaProGluAsnLeuGl  134
      |||:::||| ||||||||||||||||||||:::|||||||:||||||||
  297 TTCAGCTGTGTCCCTCGCTGCCTGGCAACGGGACAAATGGATGGAGAACCTCAA  248

  134 nArgAlaValLysProAsnLysAspAspAsnSerArgArgValAspAsnValL  151
      :|||::||||||:||||||||||:::|||||||  |:||:::::||||||
  247 AAGAACTGCAACCTAATAAGGACAATTGCCGGCGGGGGGGGAAAACGTCC  198

  151 euLysTrpIleIleGluValArgGluLeuProProLysLysArgTyr  167
      ||:::|||||:||||||:::|||||||||:::::||||||||||||
  197 TCCGCCTGTGGATCATTGAGCAAAGACCTGCCACCAAAAAGAAAATAT  148
```

```
  168 TyrCysGluCysLysAspAspAspMetLeuTyrAlaArgThrThrSerIly  184
      :::||||||||||||||||||:::|||||||||||||||||||||||
  147 TTCTGTGAGCTGTGTCTGGATGACATGCTGTACGCTCGTACCACCAGCAA  98

  184 sProArgSerAlaSerGlyAspThrValPheTrpGlyGluHisPheGluP  201
      |  ||||||   ||| ::||||||||||||||||||||||||||
   97 AACCCGCTCC........GACTGCCTGTTCTGGGGGAACACTTTGAGT  57

  201 heAsnAsnLeuProAlaValArgAlaLeuArgLeuHisLeuTyrArgAsp  217
      ||:::||||||:||:|||||::|||||||||||||||||||||||
   56 TTTCCGGCTCCGTCCATGAAGAGCATCACCATACACATCTACCGGGAT  7

  218 SerAsp  219
      |||
    6 GTGGAC  1

  seq_name: gb_est71:BE490923

  seq_documentation_block:
  LOCUS       BE490923         628 bp     mRNA            EST           31-JUL-2000
  DEFINITION  db38g08.x1 Blackshear/Soares normalized Xenopus egg library Xenopus
              laevis cDNA clone IMAGE:3300830 3' similar to TR:O95174 O95174
              NGAP..;, mRNA sequence.
  ACCESSION   BE490923
  VERSION     BE490923.1  GI:9610456
  KEYWORDS    EST.
  SOURCE      African clawed frog.
    ORGANISM  Xenopus laevis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
              Xenopodinae; Xenopus.
  REFERENCE   1  (bases 1 to 628)
  AUTHORS     Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
              , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
              ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
              Waterston,R. and Wilson,R.
  TITLE       WashU Xenopus EST project, 1999
  JOURNAL     Unpublished (1999)
  COMMENT     Other_ESTs: db38g08.y1
              Contact: Sandy Clifton, Ph.D.
              WashU Xenopus EST project, 1999
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              Library constructed by Bento Soares and M. Fatima Bonaldo
              (University of Iowa). DNA Sequencing by: Washington University
              Genome Sequencing Center
              Clone distribution: Xenopus clones from this library are available
              through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
              Seq primer: -40UP from Gibco
              High quality sequence stop: 454.
  FEATURES             Location/Qualifiers
       source          1..628
                       /organism="Xenopus laevis"
                       /db_xref="taxon:8355"
                       /clone="IMAGE:3300830"
                       /clone_lib="Blackshear/Soares normalized Xenopus egg
                       library"
                       /sex="female"
                       /tissue_type="unfertilized egg"
                       /cell_type="unfertilized egg"
                       /dev_stage="unfertilized egg"
                       /lab_host="DH10B"
                       /note="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI;
                       PolyA-selected mRNA was prepared from unfertilized Xenopus
                       laevis eggs. The library was constructed in the vector
                       pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
                       Soares, M.B. 'Normalization and subtraction: two
                       approaches to facilitate gene discovery', Genome Research
                       6:791-806, 1996. The first strand synthesis used a
```

OM of: US-09-294-298-6 to: N_Geneseq_36:*   out_format : pfs

Date: Jan 19, 2001 2:58 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US09294298/runat_17012001_125502_27790/app_query.fasta_1.4359
-DB=N_Geneseq_36 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09294298_@CGN1_1_381 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-294-298-6
Query length: 1325
Database: N_Geneseq_36:*
Database sequences: 480022
Database length: 187831343
Search time (sec): 314.010000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
/SIDS6/gcgdata/geneseq/geneseq/NA1992.DAT:Q26078	+	437.50	304.25	3.7e-09	4307	!
/SIDS6/gcgdata/geneseq/geneseq/NA1991.DAT:Q10988	+	435.50	302.86	4.4e-09	2647	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T76739	-	429.50	302.28	4.8e-09	2647	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:Z23430	+	283.00	196.84	0.0036	4279	!
/SIDS6/gcgdata/geneseq/geneseq/NA1998.DAT:V62176	-	282.50	192.03	0.0057	117213	!
/SIDS6/gcgdata/geneseq/geneseq/NA2000.DAT:Z58980	+	281.50	191.21	0.0073	7962	!
/SIDS6/gcgdata/geneseq/geneseq/NA2000.DAT:Z87283	+	279.50	184.72	0.0168	15872	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T68715	+	274.50	181.24	0.0263	15872	!
/SIDS6/gcgdata/geneseq/geneseq/NA1995.DAT:Q86927	+	273.00	188.81	0.0099	4947	!
/SIDS6/gcgdata/geneseq/geneseq/NA1995.DAT:Q75977	+	266.00	182.34	0.0228	6151	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:X56436	-	260.50	180.66	0.0283	4598	!
/SIDS6/gcgdata/geneseq/geneseq/NA1998.DAT:V62154	-	259.50	168.73	0.1308	21034	!
/SIDS6/gcgdata/geneseq/geneseq/NA2000.DAT:Z87318	+	254.00	160.77	0.3628	36778	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:Q29263	+	253.00	174.21	0.0648	5438	!
/SIDS6/gcgdata/geneseq/geneseq/NA1996.DAT:T06978	+	252.00	176.49	0.0483	3632	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:X21373	+	252.00	176.49	0.0483	3632	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:V81748	+	252.00	174.98	0.0586	4456	!
/SIDS6/gcgdata/geneseq/geneseq/NA1995.DAT:Q80016	+	251.00	185.23	0.0158	1014	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:X36583	+	251.00	170.04	0.1106	7918	!
/SIDS6/gcgdata/geneseq/geneseq/NA2000.DAT:Q20602	+	251.00	169.14	0.1240	8937	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T46941	+	251.00	167.81	0.1472	10706	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T80415	+	251.00	165.83	0.1896	13987	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T78508	-	251.00	157.30	0.5665	44377	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T80414	-	251.00	157.30	0.5665	44377	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:X58463	+	250.50	173.30	0.0727	4857	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T86704	+	250.00	176.16	0.0504	3147	!
/SIDS6/gcgdata/geneseq/geneseq/NA1995.DAT:Q96261	+	249.00	175.29	0.0563	3221	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T89345	+	247.50	172.84	0.0772	3901	!
/SIDS6/gcgdata/geneseq/geneseq/NA1994.DAT:Q73500	+	247.00	167.13	0.1605	8438	!
/SIDS6/gcgdata/geneseq/geneseq/NA1995.DAT:Q84659	+	247.00	167.36	0.1559	7808	!
/SIDS6/gcgdata/geneseq/geneseq/NA1991.DAT:Q13669	+	246.50	174.20	0.0649	2954	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T80413	-	246.00	154.00	0.8642	43280	!
/SIDS6/gcgdata/geneseq/geneseq/NA1993.DAT:Q50631	+	245.00	173.46	0.0712	2832	!
/SIDS6/gcgdata/geneseq/geneseq/NA1998.DAT:V50431	+	245.00	166.57	0.1724	7193	!
/SIDS6/gcgdata/geneseq/geneseq/NA1994.DAT:Q53954	+	243.00	171.12	0.0962	3221	!
/SIDS6/gcgdata/geneseq/geneseq/NA1998.DAT:V66549	+	242.50	174.11	0.0656	2051	!
/SIDS6/gcgdata/geneseq/geneseq/NA1996.DAT:T17420	+	242.50	173.90	0.0674	2111	!
/SIDS6/gcgdata/geneseq/geneseq/NA1998.DAT:V66546	+	242.50	173.90	0.0674	2111	!
/SIDS6/gcgdata/geneseq/geneseq/NA1991.DAT:Q14967	+	242.50	173.66	0.0695	2181	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:X58429	+	242.00	171.78	0.0884	2681	!
/SIDS6/gcgdata/geneseq/geneseq/NA1996.DAT:T27254	+	241.50	170.49	0.1044	3048	!
/SIDS6/gcgdata/geneseq/geneseq/NA1999.DAT:Z23429	+	241.50	169.33	0.1210	3563	!
/SIDS6/gcgdata/geneseq/geneseq/NA1998.DAT:V61588	+	241.50	169.26	0.1221	3596	!
/SIDS6/gcgdata/geneseq/geneseq/NA1997.DAT:T89346	+	241.00	169.75	0.1147	3211	!
/SIDS6/gcgdata/geneseq/geneseq/NA1994.DAT:Q73500	+	240.50	162.26	0.2996	8438	!

seq_name: /SIDS6/gcgdata/geneseq/geneseq/NA1992.DAT:Q26078

seq_documentation_block:
ID   Q26078 standard; DNA; 4307 BP.
XX
AC   Q26078;
XX
DT   04-JAN-1993 (first entry)
XX
DE   Lambda clone 101.
XX
KW   GAP6; placenta; cytoplasmic; cancer; diagnosis; ras p21; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   119..1899
FT   /*tag= a
FT   polyA_signal   3996..4000
FT   /*tag= b
XX
PN   WO9210572-A.
XX
PD   25-JUN-1992.
XX
PF   06-DEC-1991;   91WO-US09312.
XX
PR   07-DEC-1990;   90US-0624299.
XX
PA   (CETU) CETUS CORP.
XX
PI   Clark R, Halenbeck R, Koths KE, Martin G, McCormick FP;
PI   O'Rourke EC, Rubinfeld B, Wong GL;
XX
DR   WPI; 1992-234628/28.
DR   P-PSDB; R25335.
XX
PT   Native type 2 GTPase activating protein (GAP) is new - for
PT   diagnosing oncogenic RAS protein associated cancers
XX
PS   Disclosure; Fig 5; 70pp; English.
XX
CC   The sequence given is a lengthened version of the Type I GTPase
CC   activating protein (GAP) GAP6. It has additional DNA 5' to GAP6 and
CC   encodes a protein with a molecular weight of 116,000 daltons. This is
CC   similar to the molecular weight of the GAP protien purified from human
CC   placenta. GAP is thought to be a cytoplasmic protein which stimulates
CC   normal ras p21 GTPase activity but does not effect the GTPase activity
CC   associated with oncogenic mutants. This sequence or fragments derived
CC   from it can be useful as cancer diagnostics, being partic. useful to
CC   diagnose for ras p21 related cancers.
XX
SQ   Sequence 4307 BP; 1280 A; 857 C; 956 G; 1214 T; 0 other;

alignment_scores:
      Quality: 437.50   Length: 568
      Ratio: 1.326   Gaps: 22
Percent Similarity: 58.099   Percent Identity: 26.232

alignment_block:
US-09-294-298-6 x Q26078

Align seg 1/1 to: Q26078   from: 1 to: 4307

96 ValHisSerIleLeuGlyGlnGluPheCysPheGluValThrThrSe 112
   |||||:|||::::::|||::||:::||::: :||::: |||||||: :::::::::::::

```
1706 GTTCATGATAGTCTCTTTGGCAGGCCAAACTGTTTCAGATAGTAGTTCA 1755
 112 rSerGlyThrLys..........CysPheAlaCysArgSerAlaA
     ::::::               ||||||
1756 GCACTTTAGTGAAGACAATACATCTTTTACTTTGCAGGAGAAACTCCAG 1805
 124 laGluArgAspLysTrpIleGluAsnLeuGln............Arg 135
     ::||||||:|||||          |||
1806 AACAAGCAGAGGATTGGATGAAAGGTCTGCAGGCCATTTGCAATTTACGG 1855
 136 AlaValLysProAsnLysAspAsnSerArg.. ArgValAspAsnValLe 151
     ::||||:|||  ::|||:||| ||           ||
1856 AAAAGTAGTCCAGGGACACATCCAATAAACGCCTTCGTCAGGTCAGCAGCCT 1905
 151 uLysLeuTrpIleIleGluAlaArgGluLeuProProLys......LysA 166
     ::||| |||||||:::::||||| |||       |||
1906 TGTTTTACATATTGAAGAAGCCCATAAACTCCCAGTAAAACATTTTACTA 1955
 166 rgTyrTyrCysGluLeuCysLeuAspAspMetLeuTyrAlaArgThrThr 182
     ::  ||||||::::||| |||:::||
1956 ATCCATATTGTAACATCTACCTGAATAGTGTCCAAGTAGCAAAAACTCAT 2005
 183 SerLysProArgSerAlaSerGlyAspThrValPheTrpGlyLeuHisPh 199
     :::|||||||  :::||:||| ||||:
2006 GCAAGG..........GAAGGGCAAAACCCAGTATGTCAGAAGAGTT 2043
 199 eGluPheAsnAsnLeuProAla.. ValArgAlaLeuArgLeuHisLeu 215
     |   |||:::|| |||:|||:::
2044 TGTCTTTGATGATCTTCCTCCGACATCAATAGATTTGAAATAACTCTT. 2092
 215 yrArgAspSerAspLysSerGlyLysLysAspLysGlyTyrValGly 231
     ::::||| ||||||:::||| :::
2093 AGTATAAAACAAAGACAAAGATCCTGATAATCCTA 2131
 232 LeuValThrValProValAlaThrLeuAlaGlyLyrArgHisPheThrGluGl 248
     :::|||||||  :||| |||:||
2132 TTTATGCGCTGCCAGTTGAGCCGATTACAGAAAGGGCATGCCACAGATGA 2181
 248 nTrpTyrProValThr..... LeuProThrGlySerGlyGlySerGlyG 263
     :|||      :|||:||
2182 ATGGTTTCTGCTGCTCCAGCTCCCATATACCATTA 2212
 263 lyMetGlySerGlyGlyGlyGlyGlySerGlyGlyGlySerGlyGlyLys 279
     ...............
2212 ..............
2213 ...AAAGGTATTGAACCAGGGTCCCTGCCGTGTTCGAGCACGATACTCTAT 2259
 280 GlyLysGlyGlyCysPro...AlaValArgLeuLysAlaAlaArgTyrGlnTh 295
     :::||:||||||:::||||||||||
2260 GGAAAAATCATGCCAAGAGAAGACTACAGAGAATTAAAGAGCTTATAC 2309
 295 rMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyrValT 312
     :|||||:||| ||| |||||||:::
2310 TGCAAAAGGAACTTCATGTAGTCTAGTGCTATGCTTTATCA........CATGTA 2350
 312 hrAsnHisTyrArgMetTyrCysAlaValLeuGluProAlaLeuAsnVal 328
     ::||| ::|||
2351 TGTGGACAAGACCGAACCGACACTCGGCCAGCATCCTACTGGAGGATTTTCT 2400
 329 LysGlyLysGluGlu.....ValAlaSerAlaLeuValHisIleLeuGl 343
     ||:::::|  :|||||| |||:||:
2401 TCACGAAAAGCTTGAATCGTTGTTCTTATGCACACTAAATGACAGAGAAA 2450
 343 nSerThrGlyLysAlaLysAspPheLeuSerAspMetAlaMetSerGluV 360
     :|||||:::  |||||
2451 TAAGC.....ATGGAAGATGAGCCACTACCCTATTTCGAGCCACAACA 2494
 360 aLAspArgPheMetGluArgGlu...HisLeuIlePheArgGluAsnThr 375
     ||||| |||||| |||
2495 CTTGCCAGCACCTTGATGGAGCAGTATACAGAACCACCTGCTACACAGTT 2544
 376 LeuAlaThrLysAlaIleGluGluThrTyrMetArgLeuGlnLysLysTy 392
```

```
 392 rLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluG 409
     ||||:::  ||  ||||||||||
2545 TGTTCATCATGCTTTGAAGACTCTATTTTAAAGATAATGGAAAGCAAGC 2594
 409 luAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaGlu 425
     ||| :::  :::::
2595 AGTCTTGTGATGAAGTCCATCAAAGTTAGAAAAAAATGAAGATGTGAAC 2644
 426 HisGlnAlaAsn......LeuArgMetCysCysGluLeuAlaLeuCysLy 440
     |||||:          :|||||||
2645 ACTAATTTAACACACCTATTGAACATACTTTCAGAGCTT...GTGGAGAA 2691
 440 sValValAsnSerHisCysValPheProArgGluLeuLysGluValPheA 457
     |::: :::::||||| |||:||
2692 AATATTCATGGCTTCAGAAATACTTCCACCGACATTGAGATATATTTACTG 2741
 457 laSerTrpArgLeuArgCysAlaGluArgGly..............
     |::                  :::...TGTTTACAGAAATCTGTTCAGCATAAGTGGCCT 2776
2742 GG........
 468 ...ArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuAr 483
     :::  ||||||||:::|||||||
2777 ACAAATACCACCATGAGAACAAGAGTTGTTAGTGGTTTTGTTTTCTTCG 2826
 483 gPheLeuCysSerProAlaIleMetSerProSerLeuPheGlyLeuMetGlnG 500
     ::: |:::||||| :::|||
2827 ACTCATCTGTCGCTGCCATCCTGAATCCACGGATGTTCAATATCATCTCAG 2876
 500 luTyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysVal 516
     ::: :::||||||| |||||
2877 ATTCTCCATCTCCATTGCTGCAAGAACACTGATATTAGTGGCTAAATCT 2926
 517 IleGlnAsnLeuAlaAsnPheSerLysPheThrSerGluAspPheLe 533
     :::|||||||||||||:::  ::|||  ::|||
2927 GTGCAGAACTGCAAATTGTGTGGAGCTAAGGAGCCCTACAT 2976
 533 uGlyPheMetAsnGluPheLeuLeuGluLeuGluTrpGlySerMetGlnGlnP 550
     ::: |||  :||||||||||
2977 GGAAGGTGTCAATCCATTCATCAAAAGCAACAAACATCGTATGATCATGT 3026
 550 heLeuTyrGluIleSerAsnLeuAspThrLeuThrAsnSerSerSerPhe 566
     |||||  :::||||: ::||
3027 TTTTAGATGAACTTGGGAATGTACCTGAACTTCCGGACACTACAGAG... 3073
 567 GluGlyTyrIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLe 583
     |||||  ||||||||||||||||
3074 CATTCTAGAACGGACCTGTCCCGTGATTAGCAGCATTGCATGAGATTTG 3123
 583 u.........TrpGluValProGlnLeuSerLysGlu... 593
     :::           ||| |||
3124 CGTGGCTCATTCAGTGAACGTTCAGTAATGAGCGTGGTGCAC 3173
 594 ......AlaLeuLysLysGlyProLeuProArgLeuLeuSerAsp 607
     :::||||  :|||||||
3174 AGCAGCACGTATTGAAAAAGCTTCTGGCTATACAGAACTGCTTCAACAA 3223
 608 IleSerThrAlaLeuArgAsnProAsn...IleGlnArgGlnProSerAr 623
     |||           |||
3224 AAACAAAACCAGTATACAAAAACCAATGATGTCAGGTAGCAGCCTTCGCC 3273
 623 gGln 624
     |||
3274 CCAG 3277
```

QM of: US-09-294-298-6 to: Issued_Patents_NA:*   out_format : pfs

Date: Jan 19, 2001 7:58 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlp
 -Q=/cgn2_1/USPTO_spool/US0924298/runat.17012001_125501_27768/app_query.fasta_1.4359
 -DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US0924298 @CGN1_1_163 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:
Query: US-09-294-298-6
Query length: 1325
Database: Issued_Patents_NA:*
Database sequences: 280836
Database length: 80580151
Search time (sec): 156.930000

score_list:
Sequence   Strd Orig   ZScore   EScore Len ! Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-24 + 437.50   420.83   6.9e-16   3456 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-7 + 437.50   418.67   9.1e-16   4307 !
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-909-954-3   + 371.00   355.22   3.1e-12   3117 !
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-909-954-1   + 360.00   344.10   1.3e-11   3148 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-557-139-1   + 276.00   253.38   1.5e-06   6152 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-781-802-7   - 253.00   236.91   1.2e-05   3147 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-551-531-1   + 251.00   228.49   3.6e-05   6060 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-449-933-1   + 251.00   224.68   5.5e-05   8937 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-411-389-1   + 251.00   222.91   7.3e-05   10706 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-13 - 251.00   220.29   0.0001   13987 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-198-1   - 251.00   208.98   0.0004   44377 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-198-1   + 251.00   208.98   0.0004   44377 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-574-959A-6 + 248.50   230.30   2.8e-05   3901 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-945-283-1   - 247.50   221.74   8.4e-05   8438 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-149-097D-22 + 247.00   222.00   8.2e-05   7808 !
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-949-386-22   + 247.00   222.00   8.2e-05   7808 !
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-450-562-22   + 247.00   222.00   8.2e-05   7808 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-1 - 246.00   204.21   0.0008   43280 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-779-801-3   + 243.50   225.99   4.9e-05   3632 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-343-785-1   + 242.50   230.59   2.7e-05   2051 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-462-221-7   + 242.50   230.59   2.7e-05   2051 !
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-946-458-7   + 242.50   230.59   2.7e-05   2051 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-343-785-1   - 242.50   230.31   2.8e-05   2111 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-462-221-1   + 242.50   230.31   2.8e-05   2111 !
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-946-458-1   + 242.50   230.31   2.8e-05   2111 !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-14418-1 + 242.00 225.69   5.1e-05   3211 !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-15327-1 + 241.50 225.70   5.1e-05   3048 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-779-801-5   + 241.50   225.70   5.1e-05   3048 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-945-283-1   + 240.50   224.08   6.3e-05   3596 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-258-261B-6 + 240.50   214.72   0.0002   8438 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-456-837-6   + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-342-6   + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-646A-6 + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-076A-6 + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-4 + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-335A-6 + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-729-214-6   + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-028-934-6   + 240.50   202.64   0.0010   28958 !
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-1 + 240.50   197.41   0.0019   49377 !

/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-882-292-1   + 239.00   223.93   6.4e-05   2650 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-331-644-1   + 239.00   223.93   6.4e-05   2650 !
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-04102-1 + 239.00 223.93   6.4e-05   28 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-779-801-4   + 239.00   221.48   8.7e-05   3632 !
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-7 + 238.50   196.45   0.0022   44377 !

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-24

seq_documentation_block:
; Sequence 24, Application US/0819068?B
; Patent No. 5760203
; GENERAL INFORMATION:
; APPLICANT: Wong, Gail L.
; APPLICANT: Martin, George
; APPLICANT: McCormick, Francis P.
; APPLICANT: Rubinfeld, Bonnee
; APPLICANT: O'Rourke, Edward C.
; APPLICANT: Clark, Robin
; TITLE OF INVENTION: GAP Gene Sequences
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;   STREET: 6300 Sears Tower, 233 South Wacker Drive
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: United States of America
;   ZIP: 60606-6402
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/190,687B
;   FILING DATE: 02-FEB-1994
;   CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/774,644
;   FILING DATE: 11-OCT-1991
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/260,807
;   FILING DATE: 21-OCT-1988
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 07/230,761
;   FILING DATE: 10-AUG-1988
; ATTORNEY/AGENT INFORMATION:
;   NAME: Gass, David A.
;   REGISTRATION NUMBER: 38,153
;   REFERENCE/DOCKET NUMBER: 27527/31898
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312/474-6300
;   TELEFAX: 312/474-0448
;   TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 24:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 3456 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 100..2709
US-08-190-687B-24

alignment_scores:
                 Quality: 437.50   Length: 568
                 Ratio: 1.326   Gaps: 22
Percent Similarity: 58.099 Percent Identity: 26.232

alignment_block:
US-09-294-298-6 x US-08-190-687B-24   ..

Align seg 1/1   to: US-08-190-687B-24   from: 1   to: 3456

```
  96 ValHisSerIleLeuGlyGlnGluPheCysPheGluValThrThrSe 112
     |||||||::|||:::::|||:::||:|||||||
1156 GTTCATGATAGTCTCTTTGGCAGGCAAACTGTTTTCAGATAGTAGTTCA 1205

 112 rSerGlyThrLys..........CysPheAlaCysArgSerAlaA 124
     :  ::::::|           ||||||
1206 GCACTTTAGTGAAGAACATTACATCTTTTACTTTGCAGGAGAAACTCCAG 1255

 124 laGluArgAspLysTrpIleGluAsnLeuGln..............Arg 135
     ::::||||||::||:::::|||||||                    |||
1256 AACAAGACGAGGATTGGATGAAAGGTCTGCAGGCATTTGCAATTACGG 1305

 136 AlaValLysProAsnLysAspAsnSerArg...ArgValAspAsnValLe 151
     :::||||::: :|||::::||||||   |||       |||   ::|
1306 AAAAGTAGTCCAGGGACACCAATAAACGCCTTCGTCAGGTCAGCAGCCT 1355

 151 uLysLeuTrpIleIleGluAlaArgGluLeuProProLys......LysA 166
     |::||||||::||:::::||||||            :         |||
1356 TGTTTTACATATTGAAGAGCCCATAAACTCCCAGTAAAACATTTTACTA 1405

 166 rgTyrTyrCysGluLeuCysLeuAspAspMetLeuTyrArgAlaArgThrThr 182
     ::  ||||||::::|||||          ::|||||  :::|||
1406 ATCCATATTGTAACACTGAATAGTGTCCAAGTAGCAAAAACTCAT 1455

 183 SerLysProArgAlaSerGlyAspThrValPheTrpValGlyHisPh 199
     :::::||           |||::::::::||||||
1456 GAAGGGCCAAAACCAGTATGGTCAGAAGAGTT 1493

 199 eGluPheAsnAsnLeuProAla...ValArgAlaLeuArgLeuHisLeuT 215
     ||:::::|||||||:::::|||   ::|::||:::::||||||
1494 TGTCTTTGATGATCTTCCTCCTGACAATCAATAGATTTGAAATAACTCTT. 1542

 215 yrArgAspSerAspLysLysArgLysLysAspLysLysAlaGlyTyrTyrValGly 231
     ::|||||||||                  :::
1543 ..........AGTAATAAAACAAAGAAAAGCAAAGATCCTGATATCTTA 1581

 232 LeuValThrValProValAlaThrLeuAlaGlyArgHisPheThrGluGl 248
     :::::::|||::::|||||||:::::||||||
1582 TTTATGCCTGCCAGTTGAGCCGATTACAGAAAGGGCATGCCACAGATGA 1631

 248 nTrpTyrProValThr.....LeuProThrGlySerGlyGlySerGlyG 263
     |::::::|||||           ::|||||
1632 ATGGTTTCTGCCTCAGCTCCCATATACCATTA.............. 1662

 263 lyMetGlySerGlyGlyGlyGlyGlySerGlyGlyGlySerGlyGlyLys 279
1662 ................................................. 1662

 280 GlyLysGlyGlyCysPro...AlaValArgLeuLysAlaArgTyrGlnTh 295
     ::::||           |||   :::::::|||||||
1663 ..AAAGGTATTGAACCAGGGTCCCGTGTTCGAGCACGGATACTCTAT 1709

 295 rMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyrValT 312
     :::|||   |||:::::|||:::::||||||
1710 GGAAAAAATCATGCCAGAAGAGACTACAGTGAATTTAAAGAGCTTATAC 1759

 312 hrAsnHisTyrArgMetLeuCysAlaValLeuGluProAlaLeuAsnVal 328
     ::::|||||           :::||||||
1760 TGCAAAGGAACTTCATGTAGTCTATGCTTTATCA........CATGTA 1800

 329 LysGlyLysGluGlu......ValAlaSerAlaLeuValHisIleLeuGl 343
     :::|||      :::         :::||||||
1801 TGTGGACAAGACCGAACACTGTGTGTTAGGATTTTTCT 1850

 343 nSerThrGlyLysAspPheLeuSerAspMetAlaMetSerGluV 360
     |::::::|||   :::||||||
1851 TCACGAAAAGCTTGAATCGTTCTTATGCACACATAAAATGACACAGAAA 1900

 360 aIAspArgPheMetGluArgGlu...HisLeuIlePheArgGluAsnThr 375
     ::::::|||||           :::|||||||
1901 TAAGC......ATGGAAGAGATGAAGCCACTACCCTATTCGAGCCAACAACA 1944

 376 LeuAlaThrLysIleGluGlyLeuTyrMetArgLeuIleGlyValIleGlnLysTy 392
     |||::::::|||:::::||||||:::::|||
1945 CTTGCAAGCACCTTGATGGAGCAGTATATGAAAGCCACTGCTACACAGTT 1994

 392 rLeuLysAspAlaIleGlyLeuPheIleGluArgAlaLeuTyrGluSerGluG 409
     |:::::|||   :::||||||
1995 TGTTCATCATGCTTTGAAAGACTCTATTTAAAGATAATGGAAAGCAAGC 2044

 409 luAsnCysGluValAspProIleLysCysThrAlaSerLeuAlaGlu 425
     |::::::|||||           :::||||||
2045 AGTCTTGTGAGTTAAGTCCATCAAAGTTAGAAAAAAATGAAGATGTGAAC 2094

 426 HisGlnAlaAsn......LeuArgMetCysCysGluLeuAlaLeuCysLy 440
     :::|||         |||   :::::::||||||
2095 ACTAATTTAACACACTATTGAACATACTTTCAGAGCTT...GTGGAGAA 2141

 440 sValValAsnSerHisCysValPheProArgGluLeuLysGluValPheA 457
     ::::::|||||           :::|||||||
2142 AATATTCATGGCTTCAGAAATACTTCCACCGACATTGAGATATATTTATG 2191

 457 laSerTrpArgLeuArgCysAlaArgGly.................... 467
     :::|||                     ::
2192 GG.............TGTTTACAGAAATCTGTTCAGCATAAGTGGCCT 2226

 468 ..ArgGluAspIleAlaAspArgGluIleSerAlaSerLeuPheLeuAr 483
        |||   :::::::||||||      :::|||
2227 ACAAATACCACCATGAGAACAAGAGTTGTAGTGGTTTGTTTTTTCTTCG 2276

 483 gPheLeuCysProAlaIlemetSerProSerLeuPheGlyLeuMetGlnG 500
     |::::::||||||||||::::::||||||
2277 ACTCACTGTCCTGCCATCCTGAATCCACGATGTCAAATATCACTCAG 2326

 500 luTyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysVal 516
     :::||||||           :::|||||||
2327 ATTCTCCATCCTATTGCTGCAAGAACACTGATATTAGTGGCTAAATCT 2376

 517 IleGlnAsnLeuAlaAsnPheSerLysPheThrSerGluLysAspPheLe 533
     ::::::|||:::::||||||      :::||||||
2377 GTGCAGAACTTAGCAAATCTTGTGGAATTTGGAGCTAAGGACCCTACAT 2426

 533 uGlyPheMetAsnGluPheLeuLeuGluTrpGlySerMetGlnGlnP 550
     |::::::|||||           :::||||||
2427 GGAAGGTGTCAATCCATTCATCAAAAGCAACAAACATGCTGATGATCATGT 2476

 550 heLeuTyrGluIleSerAsnLeuAspThrLeuThrAsnSerSerPhe 566
     :::|||                     :::|||
2477 TTTTAGATGAACTTGGGAATGTACCTGAACTTCCGGACACTACAGAG... 2523

 567 GluGlyTyrIleAspLeuGlyLeuArgGluLeuSerThrLeuHisAlaLeuLeu 583
     :::::::|||:::::||||||      :::|||
2524 CATTCTAGAACGACCTGTCCCGTGATTTAGCACATTGCAGCATTGAGATTTG 2573

 583 u.........TrpGluValProGlnLeuSerLysGlu.......... 593
     u           :::||||||
2574 CGTGGCGTCATTCAGATGAACTTCGAACGCTCCAGTAATGAGCGTGGTGCAC 2623

 594 ........AlaLeuLeuLysLeuGlyTyrProLeuArgLeuSerArgAsp 607
             :::||||||           :::|||||||
2624 AGCAGCACGTATTGAAAAAGCTTCTGGCCTATAAACAGAACTGCTTCAACAA 2673

 608 IleSerThrArgLeuArgAsnProAsn...IleGlnArgGlnProSerAr 623
     ::::::|||:::::||||||   :::|||
2674 AAACAAAACCACTATACAAAAACCATGATGTCAGGTAGCAGCCCTTCGCC 2723

 623 gGln 624
     |||
2724 CCAG 2727
```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-7

seq_documentation_block: